

Cloning Scheme for Generating pCMV-NS35

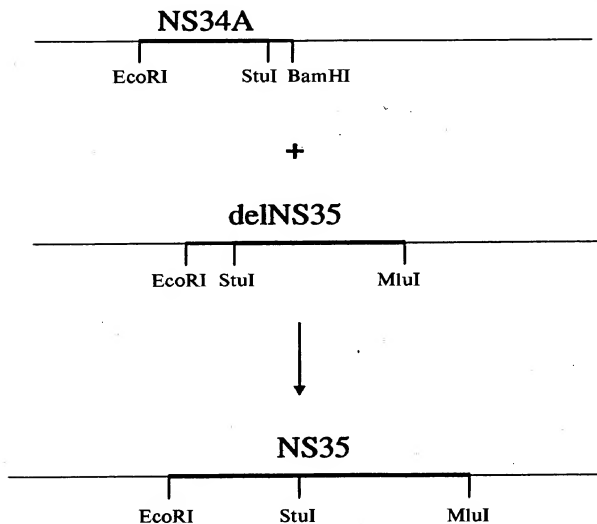


FIG. 1

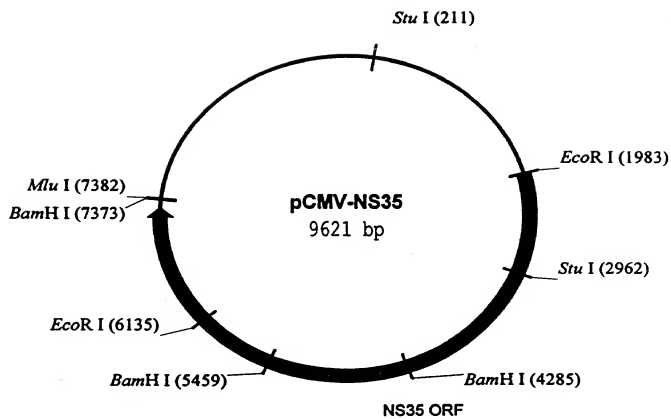


FIG. 2

pCMV-NS35

1 TCGGCGGTTT CGGTGATGAC GGTGAAGAAC TGTGACACAT CAGGTCCCG GAGACGGTCA CAGCTGTCT GTAACCGCAT
AGCGCGAAA GCCACTACTG CCACCTTTTG AGACTGTGTA CGTCGAGGCG CTCTGCCAGT GTCGAACAGA CATTGGGCTA

81 CCGCGGAGCA GACAACCGG TCAGCGCGG TCAGCGCGG TTGCGCGGTG TCAGCGCGGT CTTAACATG CCGCATCAGA
CGGCGCTCGT CTGTGCGGC AGTCGCGGC AGTCGCGGC AACCGCGCAC ACCCGCGACC GAAITGATAC CGCGTAGTCT

161 GCAGATTGTA CTGAGATGCG ACCATATGAA GCTTTTGGCA AAGCTTAGG CUTCAGAAA AGCCTCTCTCA CTACTTCTGG
CGCTAACAT GACTCTCAG TGGTATCTT CGAAAAAGCT TTTGCGATCC GGAGTTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGCGCGAGC GCGCTCGGC TCTGCATAA TAAAAAAT TACTCAGCCA TGGGCGGAG AATGGCGGA
TTATCGAGTC TCGGCGTGG CCGAGCGCGG AGACGTATTT ATTTTTTTA ATCAGTCGCT ACCCGCGCTC TTACCGCGCT

321 ACTGGCGGG GAGGAATTA TTGGCTATTG GCGATTGCA AGCTGTATC TATATCATAA TATGTACATT TATATTGGCT
TGACCGCGCC CTCCTTAAT AACCGATAAC CGGTACGTA TCGAACATAG ATATAGTATT ATACATGTAA ATATAACGA

401 CATGTCCAAT ATGACCGCCA TGTTCACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTTTCA
GTACAGGTTA TACTGGCGGT ACACTGTAA CTAATACTG ATCAATAATT ATCATTAGT AATGCCCGAG TAATCAAGTA

481 AGCCCATATA TGGAGTTCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGGTGACCG CCCAACGACC CCGCGCGAAT
TCGGGTAT ACCTCAAGC GCAATGTATT GAATGCCATT TACCGCGCGG ACCGACTGC GGGTTGCTGG GGGCGGGTAA

561 GAGTCAATA ATGACATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTATT CCTGAAAGG TAACTGCAGT TACCGACCTC ATAAATGCCA

641 AAATGCCCA CTTGGCAGTA CATCAAGTCT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACG TAAATGGCCC
TTTGACGGGT GAACCGTCAI GTAGTTTACA TAGTATACGG TTCAGCGCGG GGAATACATCC AGTTACTGCC ATTACCGGG

StuI

FIG. 3-Page 1

pCMV-NS35

721 GCTCGCAATT ATGCCAGTA CATGACCTTA CGGACTTTTC CTACTGGCA GTACATCTAC GTATTACTCA TCCTATTAC
CGGACCGTAA TACGGTCAAT CTACTGGAAT GCCCTGAAAT CATCAACCGT CATGTAGATG CATAATCACT AGGATAATG

801 CATGTGTGAT CGGTTTTGGC ACTACACCAA TGGCGTGGG TAGCGTTTG ACTCAGGGG ATTTCCAAGT CTCACCCCA
GTACCAGTAC GCCAAAACCG TCAITGGTIT ACCCGCACTT ATCCCAAC TGAGTCCCG TAAAGTTCA GAGGTGGGT

881 TTGAGTCAA TGGAGTTTTC TTTTGGACC AAAATCAAG GGACTTTTCA AATGTCTGA ATAAACCCGC CCGCTTGAG
AACTCAGTT ACCCTCAAC AAACCGTGG TTTTAGTTGC CCTGAAAGT TTTACAGCAT TATTGGGGG GGGCAACTGC

961 CAAATGGGCG GTAGCGGTGT AGCGTGGGAG GTCTATATAA CAGAGGTGC TTTAGTGAAC CGTCAGATCG CTTGGAGAG
GTTTACCGC CATCGGACA TGGCACCGTC CAGATATATT CGTCTCGAGC AATCACTTG CCACTCTAGC GGACCTCTGC

1041 CCATCCAGC TCTTTGACC TGCATAGAAG ACACCGGAC CGATCCAGCC TCOCGGGCG GGAACGGTGC ATTGGAACGC
GCTAGGTGCG ACAAAACG AGGTATCTTC TGTGGCCCTG GCTAGTCCG AGCGCCGCG CCTTGGCAG TAACCTTGG

1121 GGATTCGCG TCCAGAGT GACGTAAGTA CCGGCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGC ACGTTCTCA CTGCATTAT GCGGATATC TGAGATATCG GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCC GCTCCTTATG CTATAGTGA TGSTATAGCT TAGCTATAG GTGTGGTTTA
GACAAAAACC GAACCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT

1281 TTGACCATA TTGACACT CCCTATTGGT GACGATCTT TCCATTACTA ATCCATAA TGGTCTTTG CCACAACTAT
AACTGGTAAT AACTGGTGA GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGATGGG GTCCATTAT
GAGATAACCG ATATACGGTT ATGACAGG AAGTCTCTGA CTGTGCTGA GACATAAAA TGTCTACCC CAGGTAATA

FIG. 3-Page 2

pCMV-NS35

1441 TATTACAAA TTCACATATA CAACAAGCGG GTCCCGCGTG CCGCGAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
ATAAATGTTT AAGTGTAAT GTTGTTCCGG CAGGGGCGAC GGGGCTCAAA AATAATTTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGGTACG TCTTCCGGAC ATGGGCTCTT CTCGGGTAGG GGGCGAGCTT CCACATCCGA GCCTGTGTCC CATCCCTCCA
GAGCCCATGC ACAAGGCTG TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTTAGGCT CGGACCAGG GTAGGCAGT

1601 GGGGCTCATG GTCGTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCGAGA CTTAGGCACA GCACAATGCC CACCAACACG
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTA CCTCGGTCT GAATCCGTCT CGTTTACGG GTGTTGTGG

1681 AGTGTGCCG ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAAATGAGT CGGAGATTGG GTTCGCACCT GGACGCAGAT
TCACAGGGCG TCTTCCGGCA CCGGCATCC ATACAGAGC TTTTACTCGA GCCTCTAAC CGAGCGTGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGC AGAAGAGAT GCAGGCAGT GAGTTGTCT ATTCTGATA GAGTCAGAGG TAATCCCGT
CCTTCTGAAT TCCGTGCGG TCTTCTCTA CGTCCGTGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGCA

1841 TCCGCTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGA GTACTGTTG CTGCGCGCG CGCCACCAGA CATAATAGT
ACGCCAGCAC AATTGCCACC TCCGTCACA TCAGACTCT CATGACCAAC GACGGCGCG. GCGGTGGTCT GTATTATCGA

+2

M A A

EcoRI

1921 GACAGACTAA CAGACTGTC CTTTCCATGG GTCTTTTCTG CAGTCACCTT CGTCGACCTA AGAATTCACC ATGGCTGCAT
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGTATAAG GTGCTAGTAC TCAACCCCTC TGTGTGCA ACACTGGCT TTGTGCTTA CATGTCCAAG
TAGCTGAGT CCCGATATC CACGATCATG AGTTGGGGAG ACAACGACT TGTGACCCGA AACCAGGAAT GTACAGTTC

FIG. 3-Page 3

PCMV-NS35

| | |
|------|---|
| +2 | A H G I D P N I R T G V R T I T T G S P I T Y S T Y G |
| 2081 | GCTCATGGGA TCGATCTTAA CATCAGAGC GGGGTAGAA CAATTACGAC TGGCAGCC ATCAGTACT CCACCTACGG |
| | CGAGTACCT ACCTAGGATT GTAGTCTGG CCCACTCTT GTTAATGGTG ACCGTGGGG TAGTGCATGA GGTGGATGCC |
| +2 | K F L A D G G C S G G A Y D I I I C D E C H S T D A |
| 2161 | CAAGTTCTTT GCCAGGCGG GGTGCTGGG GGGGCTTAT GACATAATAA TTGTGACGA GTGCCACTCC ACGGATGCGA |
| | GTTCAGAGAA CCGCTGCCG CCAAGAGCC CCCCAGATA CTGTATTATT AAACACTGCT CACGGTGAGG TGGCTACGGT |
| +2 | T S I L G I G T V L D Q A E T A G A R L V V L A T A T |
| 2241 | CATCCATCTT GGGCATTGGC ACTGTCTTG ACCAAGCAGA GACTGGGGG GCGAGACTGG TTGTGCTGC CAGCGCACCC |
| | GTAGGTAGAA CCGCTAACCG TCACAGAAC TGGTTCTCT CTGACGCCCC CGCTCTGACC AACACGAGGG GTGGCGGTGG |
| +2 | P P G S V T V P H P N I E E V A L S T T G E I P F Y G |
| 2321 | CTTCGGGGCT CGGTCACTGT GCCCATCCG AACATCGAGG AGGTTGCTCT GTCCACACG GGAGAGATCC CTTTTACGG |
| | GGAGGCCGA GCCACTACA CCGGTAGGG TTGTAGTCC TCCAACGAGA CAGTGTGG CCTCTCTAGG GAAAAATGCC |
| +2 | K A I P L E V I K G G R H L I F C H S K K K C D E L |
| 2401 | CAAGCTATC CCGCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTCT GTCAATTCAA GAAGAAGTGC GAGGAACCTGG |
| | GTTCGGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA CAGTAGAAGA CAGTAAGTTT CTTCCTCAGG CTGCTTGAGC |
| +2 | A A K L V A L G I N A V A Y Y R G L D V S V I P T S G |
| 2481 | CGGCAAGCT CGTGCATTG GGCATCAATG CCGTGGCCTA CTACCGGCT CTTCGAGTCT CCGTCATCC GAGCAGCGGC |
| | GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGGCCA GAATGCACA GGCAGTAGGG CTGGTCGGCC |
| +2 | D V V V V A T D A L M T G Y T G D F D S V I D C N T C |
| 2561 | GATGTTGTGG TGGTGGAAAC CGATGCCCTC ATGACCGGCT ATACGGCA CTTCGACTGG GTGATAGCT GCAATACGTT |
| | CTACACAGC ACCACCGTTG GCTAGGGGAG TACTGGCCGA TATGGCCGT GAAGCTGAGC CACTATCTGA CGTTATGGAC |

FIG. 3-Page 4

pCMV-NS35

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 TGTACCCAG ACAGTCGATT TCAGCTTGA CCTACCTTC ACCATTGAGA CAATCAGCCT CCCCAAGAT GCTGTCTCCG
 ACAGTGGCTG TGTAGCTAA ACTCGGAAT CCGATGGAAG TGTAACTCT GTTAGTGGA GGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCATCAACG TCGGCGAGG ACTGGCAGG GGAAGCAGG CATCTACAGA TTTGTGSCAC GGGGGGAGG CCGCTCGGCG
 CGTAGTTGC ACCCCCTCC TGACCTCC CTTTCGGTCC GTAGATCTT AAACACCGTG GCGCCCTCC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTCACT CTTCCCTCT CTGTAGTGG TATGAGCAG GCTGTGTTG GTATGAGTC ACGCCGCGG AGACTACAT
 TACAAAGCTGA GCGGCGAGG GACACTCAG ATACTGGCTC CGACAGGAAC CATACTCGAG TGGGGGCGG TCTGATCTA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 StuI

 2881 TAGGCTACGA GCGTACATGA ACACCCGGG GCTTCCGGT TCGCAGGACC ATCTGAATT TTGGGAGGCG GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGCGG GAAAGGGCAC ACGGTCTCG TAGAACTTAA AACCTTCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
 StuI

 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAA CGAGAGTGG GAGAACTTC CTTACCTGCT AGGTAACAA
 CGGAGTACT ATATCTACGG GTGAAAGATA GGTCTCTTT CGTCTCACCC CTCTTGGGAG GAATGGACCA TCGCATGGTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
 3041 GGCACCGTGT GGGGTAGGG TCAAGCCCTT CCCCATCGT GGGACAGAT GTGGAAGTGT TTGATTGCG TCAAGCCAC
 CGGTGGACA CCGCATCCG AGTTGSGGGA GGGGTAGCA CCCTGTGCTA CACCTTCACA AACTAAGCG AGTTCCGGGTG

FIG. 3-Page 5

pCMV-NS35

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121 C C C A T G G G C A A C A C C C T G T A T A C A G A C T G G G C G T G T T C A A G T G A A T C A C C A C C C G A C G A C C C A G T C A C A A A T
G G A G T A C C G G T T G T G G G A G G A T A T G T C T G A C C G G C A A A G T C T T A C T T T A G T G G A C T G C G T G G G T C A G T G T T A

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201 A C A C T A T G A G A T G C A T G G C G G A C T G G A G G T C G T A C A G A C C T G G G T C G T G T G T G G C G C G G T G C T G C T G T T G
T C T A G T A C T G T A C G T A C A G C G G C T G G A C C T C C A G C A G T C T G T G G A C C C A G A G C A A C C G C C G A A C C G A G A A A C

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281 G C C G C G T A T T G C T G T C A A C A G G T C G T G T G C A T A G T G G C A G G G T G T C T T G T C G G G A A G C C G G A A T C A T A C T G A
C G C G G C A T A A C G G A C A G T T G T C C G A C G C A C A G T A T C A C C G T C C C A G C A G A A C A G A G G C C T T C G G C C G T T A G T A T G G A C T

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361 C A G G A A G T C C T C T A C C G A G A G T T C G A T G A G A G A G T G T C T C A C C A C T T A C C G T A C A T C G A G A A G G T A T G C C
G T C C C T T C A G A G A T G G C T C A A G C T A C T C T A C C T T C T A C G A G T G G T G A A T G G C A T G T A G T C G T T C C T A C T A C G

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441 T C G C C G A G A G T T C A A G A A G C C T G G C C T C T G A G A C C G G T C C C G T C A G G A G A G T T A T G C C C C T G C T C
A G C G G C T C G T C A A G T T C G T T C G G G A G C C G G A G A C C T C T G G C G A G G C A G T C C G T C C A A T A G C G G G A C A G

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521 C A G A C C A A C T G C A A A A A T C G A G A C T T C T G G C G A A G C A T A T G G A A C T T C A T C A G T G G A T A C A A C T A C T T G C G G G
G T C T G T T T G A C C G T T T T T G A C T C T G G A A G A C C G G C T T G T A T A C A C C T T G A A G T A G T C A C C C T A T G T T A T G A A C C G C C C

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 C T T G T C A A G C T G C T G T A A C C C G C C A T T G T T C A T T G A T G T G T T T A C A G T C T C T C A C A G C C C A C T A A C A C A C T A
G A A C A G T T G C A C G G A C C A T T G G G G G G T A A C G A A G T A C T A C C G A A A T G T C G G A C A G T G G T C G G G T G A T T G T G T A T

FIG. 3-Page 6

PCMV-NS35

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 - GCCAACCTT CCGTTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC CGCGCCCGG GTCCGCTAC TGCTTTTG
CGGTTTGGGA GGGAAGTTG TATAACCCG CCACCCACG ACGGGTGGAG CGCGGGGCG ACAGCGGAT ACAGAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGCTGGCT TAGTGGCGC CGCATCGC AGTGTGGAC TGTTGGAGT CCTCATAGAC ATCTTGAG GGTATGGCG
CGCGGACCGA ATCGACGGG CGGTAGCGG TCACAACCTG ACCCTTCA GGATATCTG TAGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GCGGCTGGC GGAGCTTTG TGCGATTCA GATCATGAC GGTGAGTCC CTCCAACGGA GCACCTGGTC AATCTACTGC
CCGCGACCG CCTCGAGAAC ACCGTAAGT CTAGTACTG CCACTCCAG GGAGTGGCT CTGCGACCG TTAGATGAGC

+2 P A I L S P G A L V G V V C A A I L E R H V G P G E
3921 CGGCTACTT CTCGCCGGA GCCTCGTAG TCGGCGTGT CTTGAGCA ATACTGGCC GCAGCTTG CGCGGCGGAG
CGCGGTAGGA GAGGGGCGCT CGGAGCATC AGCGGACCA GACACTGCT TATAGCGCG CCGTCAACC GCGCGCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GCGGCTGCG AGTGGATGAA CCGGCTGATA GCCTGGCT CGCGGGGAA CCATGTTCC CCACGCACT ACGTCCCGA
CCCGTCAAG TCACCTACTT GCGCGACTAT CGGAAGCGA GGGCGCGCTT GGTACAAAG GGTGCGTCA TGCAGGCGCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGGATGCA CTGCGCGG TCATGCCAT ACTCAGCAGC CTCAGTGTAA CCAGCTCCT GAGGCGACTG CACCGCGG
CTCGCTACCT CGAGGGGCG AGTCAGGGA TGAAGTCTG GAGTGACATT GGTGCGGGA CTCGCTGAC GTGCTCACT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
4161 TAAGTCGGA GTGTACCAT CCAATGTCG GTTCGTGCT AAGGACATC TGGAGTGA TATGGAGGT GTTGAGCGAC
ATTGCGCCT CACATGCTGA GTACGAGGC CAAGGACCA TTCCCTGAT ACCCTGACT ATACGTCCA CAATCTGCTG

pCMV-NS35

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
BamHI

4241 TTAAAGACCT GGTAAAGC TAAGTCATG CCACAGCTCG CTGGATCCC CTTTGTGCC TGCACGGCG GGTATAAGG
AAATCTGGA CGATTTTCG ATTCGAGTAC GGTCTGAGG GACCTAGG GAACACAGG ACCTGCGCG CCATATCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T
4321 GGTCTGGCGA GGGGACGGCA TATGACACAC TCGTCCAC TGTGAGCTG AGATCACTG ACATGTCAA AACGGAGGA
CCACAGCGCT CCCCTGCCGT ACTACGTGTG AGCGACGGT ACACCTCGAC TCTAGTACC TTACAGTTT TTGCCCTGT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C
4401 CTGAGATCGT CGGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGGACTTC CCCATTATG CCTACACCAC GGGCCCCGT
ACTCCTAGCA GCGAGGATCC TGGAGCTCCT TGTACACCTC ACCCTGGAAG GGTTAATTAC GGATGTGGT CCGCGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
4481 ACCCCCTTC CTGCGCGGAA CTACAGCTTC CGCTATGGA GGGTGTCTG AGAGGAATAC GTGGAGATAA GGCAGGTGG
TGGGGGAAG GACGGGGCTT GATGTGCAAG GCGGATACCT CCACAGAGG TCTCCTTATG CAGCTCTATT CCGTCCACC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T
4561 GCACTTCAC TAGCTGACG GTATGACTAC TGACAATCTT AAATGCCCTT GCGAGTCCC ATGCCCCGAA TTTTTCACAG
CCTGAAGGTG ATGCACTGCC CATACTGAT ACTCTTAGAA TTTACGGGA CCGTCCAGG TAGCGGGCTT AAAAGTGT

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G
4641 AATTCGAGG GGTGCGCTA CATAGTTTG CGCCCCCTG CAAGCCCTG CTGGGGAGG AGTATCATT CAGATAGGA
TTAAGCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGAC GTTCCGGAAC CAGCGCTCC TCCATAGTAA GTCTCATCT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCACGAAT ACCGGTAGG GTCCCAATTA CTTGCGAGC CCGAACCGGA CGTGGCGGTG TTGACCTCCA TGTCTACTGA
GAGGTGCTTA TGGGCCATCC CAGGGTTAAT GGAACGCTCG GGGTTGGCTT GCACCGGGAC AACTGCAGGT ACCAGTACT

FIG. 3-Page 8

pCMV-NS35

-42 P S H I T A E A A G R R L A R G S P P S V A S S A
4801 TCCCTCCCAT ATACAGCAT AGCGCGCG GCGAAGTTG CACCCCTC TTGGCCAGC TCCTGGCTA
AGGAGGTA TATTGTGTC TCGCCGCGC CGCTCCAC CGCTCCCTA GTGGGGGAG ACACGGTCG AGAGCCAT

+2 S O L S A P S L K A T C T A N H D S P D A E L I E A N
4861 GCCAGTATC CGCTCCAT CTCAAGCAA CTTGACCGC TAACATGAC TCCCTGTATG CTGACTCAT AGAGCCCAAC
CGGTCGATAG CCGAGGTAGA GAGTCCGTT GAACGTGGCG ATTGTACTG AGGGACTAC GACTCGAGTA TCTCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGGA GCGAGGAT GGGCGCAAC ATCAGCGG TTGAGTCAGA AAACAACTG GTGATTCTG ACTCCTCGA
GAGGATACCT CCGTCTCTA CCGCGCGTTG TAGTGTC CACTAGTCT TTGTTTCAC CACTAAGACC TGAGAACT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCCTTTGTG CCGAGGAGG ACAGCGGA GATCTCCTA CCGCAGAA TTCTGGGA GTCTGGAGA TTGGCCAGG
AGGGAACAC CGCTCTCT TGTCTGCT CTAGAGCAT GGGCTCTTT AGACGCTT CAGAGCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCCTGCCCT TTGGCGCG CCGACTATA ACCCCCGCT AGTGGAGC TCGAAGAG CCGACTAGA ACCACTGTG
GGGACGGCA AACCGCGC GGCTGAT TGGGGGGGA TCACCTCTGC ACCTTTTCG GGTGATGCT TGTGGACAC

+2 V H G C P L P P P K S P P V P P R K K R T V V L T E
5201 GTCCATGCT TTGGCGTCC ACCTCAAG TCCCTCTG TCCCTCGC TCGAAGAG CCGAGGTGG TCCTACTGA
CAGTAGCGA CCGCGAAG TGGAGTTTC AGGGAGGAC ACAGAGCGG AGCTTCTTC GCGTCCACC AGAGTGACT

+2 S T L S T A L A E I A T R S F G S S S T S G I T G D
5281 ATCAACCTA TCTACTGCT TTGGCGACT CGCAGCAGA AGCTTTGCA GTCTCTAAC TTCCGGCAT ACAGCGGCA
TAGTTGGAT AGATAGCGA ACCGGCTGA CCGGTGCT TCGAAACGT CGAGAGTTG AAGGCCGTAA TGCCGCTGT

FIG. 3-Page 9

pCMV-NS35

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATACGACAAC ATCTCTGAG CCGCGCCCTT CTGGCTGCC CCGGACTCC GACGCTGACT CCTATTCCTC CATGCCGCC
TAATCTGTG TAGGAGACTC GGGGGGGGAA GACCGACGGG GGGGCTGAGG CTGGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V

BamHI

5441 CTGGAGGGG AGCCTGGGGA TCGGATCTT AGCGAGGGT CATGCTCAAC GGTCACTAGT GAGGCCAACG CGGAGGATGT
GACCTCCGCC TGGGACCCCT AGGCTAGAA TCGCTGCCA GTACCACTTG CCACTCATCA CTCGGGTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGTGCTGC TCAATGTCTT ACTCTGGAC AGCGCACTC GTACCCCGT CGCGCGGGA AGACAGAAA CTGCCCATCA
GCACACGAGC AGTTACAGAA TCAGAACCTG TCGCGTGAG CAGTGGGCA CGCGCGCCT TCTTGCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCAGTAAG CAACCTGTG CTAGCTCAC ACAATTGGT GTATTCAAG ACCTCAGCA GTGCTTGCCA AAGCGAAG
TAGCTGATTC GTTAGCAAC CATGCACTGG TGTTAAACCA CATAAGTGG TGGAGTGGT CACGAACGGT TTGCGTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
5681 ATTCAGCAT TTGACAGT GCAAGTCTG GACAGCCAT ACTCAAGGAG GTTAAAGCAG CGGCTCAA
TTTCAGTGA AACTCTCTGA CGTTCAAGAC CTGTCGGTAA TGTCCTGCA TGAGTCTCTC CAATTCTGC CGCGCAGTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AAGTAAGCT AACTTGCTAT CGGTAGAGGA AGCTTGACG CTGAGGCC CACACTCAG CAATCAAG TTGCTTATG
TCACCTCGGA TTGAACGATA GGCATCTCT GACTCGGGG GTGTGACTGG GTTAGCTTC AAACCAATAC

FIG. 3-Page 10

pCMV-NS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGGAAAGA GTCCGTTGC CATGCCAGAA AGCCGTAAC CCACATCAAC TCCGTGTGA ATCAGCTTCT GGAAGCAAT
GCCCTTTTCT GCAGGCAAG GTACGGCTTT TCCGGCATTG GGTGTAGTT AGGCACGCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
5921 GTAACACCA TAGACACTAC CATCATGGCT AAGACGAGG TTTTCTGGGT TCAGCCTGAG AAGGGGGTTC GTAAGCCAGC
CAATTGCGTT ATCTGTGATG GTAGTACCGA TTCTTGCTC AAAAGACGA AGTCGGACTC TTCCCCCCAG CATTCGGTGC

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCGTCTCAT GTGTTCGCG ATCTGGCGT CGCGGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGTCCCT
AGCAGAGTAG CACAAGGGG TAGACCGCA CGCGCACAG CTTTCTACC GAAACATGCT GCACCAATCT TTGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

6081 TGGCCGTGAT GGAAGTCC TAGGATTCC AATACTCAC AGGACAGCG GTTGAATTCC TCGTCAAGC GTGGAAGTCC
ACCGGCACTA CCCTTCGAG ATGCTAAG TTATGATGG TCCTGTCCG CAACTTAAG AGCAGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAACCC CAATGGGTT CTGCTATGAT ACCGCTGCT TTGACTCCAC AGTCACTGAG AGGCATCC GTACGGAGGA
TTCTTTTGGG GTTACCCAA GAGCACTA TGGCCAGCA AACTGAGTG TCACTGACTC TCCTCTAGG CATGCTCTCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGCAATCAC CAATGTTGT ACCTGCACC CCAAGCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGG
CGGTTAGTG GTTACAACAC TGAAGTGGG GTTTCGGGG CACCGTAGT TCAGGGAATG GCTCTCCGAA ATACAACCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTTTAC CAATTCAAG GGGAGAAT CGGCTATCG CAGGTGCGC GCAGGGGCG TACTGACAAC TAGTGTGTT
CGGAGAAATG GTTAAATTCC CCCCTCTTGA CGCGATAC GTCCAGGCG CGTCTCCGCG ATGACTTTG ATGACACCA

FIG. 3-Page 11

pCMV-NS35

+2 N T L T C Y I K A R A A G L Q D C T M L V C G
6401 AACACCTCA CTTCGTACAT CAAGGCCGG GCAGCTCTC GAGCGAGG GTCCAGGAC TCACACATGC TCGTGTGTGG
TTGTGGAGT GAACGATGA GTTCGGGCC CGTCGGACG CTGGGGCTCC CGAGCTCTG AGTGTGTAGC AGCAGACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481 CGAGCTTGA GTGTATTCT GTGAAGGCG GGGGTCCAG GAGGAGGGG CGAGCTCAG AGCTTTACG GAGGCTATGA
GCTCTGAAT CAGCAATGA CACTTTCGG CCCCCAGTC CTCTCGGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
6561 CCAGGTACTC GGCCTCCCT GGGGAGCGC CAGAGCGCC CAGACAGA ATACGACTG GAGCTCATAA CATCATGCTC CTCCAAGCTG
GCTCCATGAG CGGGGGGGGA CCCCCTGGGG GTCTTGCTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGTTGAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641 TCAGTGGCC ACAGCGCGC TGAAGAGG GTCTACTAC TCACCGTGA CCTACAACC CCCCCTCGGA GAGTGGCTG
AGTCAGGGG TGCTGGCGG ACCTTCTCC CAGATGATG AGTGGGACT GGGATGTTGG GGGGAGCGCT CTCGACGAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721 GGAGACAGA AGACACATC CAGTCAATC CTGGCTAGC AACATAACA TGTTCGCC CACACTGTG GCGAGGATGA
CCTCTGCT TCTGTGTAG GTCAATTAG GACGATCCG TTGTATTAGT ACAAGCGGG GTCTGACACC CGCTCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
6801 TACTGTAGC CATTCTTT AGGTCTCTTA TAGCCAGGA CCAGCTTGA CAGGCCCTCG ATTGGAGAT CTACGGGGC
ATGACTACTG GTTAAGAA TCSCAGGAAT ATCGTCCCT GTCTGAACT GTCCGGGAGC TAACGCTCTA GATGCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881 TGCTACTCCA TAGAACCTT GATCTACT CCAATCATT AAAGATCCA TGCCCTCAGC GGATTTTAC TCACAGATTA
ACGATGAGT ATCTTGCTA CCTAGATGA GTTACTAAG TTTCTGAGT ACCGAGTGC CGTAAAGTG AGGTGCTCAAT

FIG. 3-Page 12

pCMV-NS35

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------------|------------|-----------|------------|------------|------------|-----------|------------|------------|------------|------------|------------|------------|-----------|------------|------------|---|---|---|---|---|---|---|---|---|---|
| +2 | S | P | G | E | I | N | R | V | A | A | C | L | R | K | L | G | V | P | P | L | R | A | W | R | H | R |
| 6961 | CTCTCCAGGT | GAAATCAATA | GGGTGCGCC | ATGCCTCAGA | AAATCTGGGG | TACGCGCCTT | GCAGCTTGG | AGACACCGGC | GAGAGGTCCA | CTTTAGTTAT | CCCAACGGGG | TACGGAGTCT | TTTGAACCCC | ATGGCGGAA | CGCTCGAACC | TCTGTGGGCC | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------------|-----------|-----------|-----------|------------|------------|-----------|------------|-----------|-----------|-----------|-----------|------------|------------|------------|-----------|---|---|---|---|---|---|---|---|---|---|---|
| +2 | A | R | S | V | R | A | R | L | L | A | R | G | G | R | A | A | I | C | G | K | Y | L | F | N | W | A | V |
| 7041 | CCCGAGCGGT | CCGGGTAGG | CTTCTGGCA | GAGAGGCAG | GGTGGCAATA | TGTGGCAAGT | ACCTCTCAA | CTGGGCAGTA | GGGCTCGCA | GGCGCATCC | GAAGACGGT | CTCTCCGTC | CCGACGGTAT | ACACCGTTCA | TGGAGAAGTT | GACCGTCAT | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------------|------------|-----------|----------|-----------|-----------|------------|------------|-----------|------------|-----------|-----------|------------|------------|------------|------------|---|---|---|---|---|---|---|---|---|---|---|
| +2 | R | T | K | L | K | L | T | P | I | A | A | A | G | Q | L | D | L | S | G | W | F | T | A | G | Y | S | G |
| 7121 | AGAACAAAGC | TCAAACCTAC | TCCAATAGG | GCGCTGGC | AGGTGACTT | GTCGGCTGG | TTCAGCGGTG | GTCACAGCGG | TCTTGTTCG | AGTTTGAGTG | AGGTTATCG | CGGCGACCG | TGCACCTGAA | CAGGCGGACC | AAGTGGCGAC | CGATGTGCCC | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|-----------|-----------|-----------|-----------|------------|------------|-----------|------------|------------|------------|------------|-----------|------------|-------------|------------|-----------|---|---|---|---|---|---|---|---|---|---|
| +2 | G | D | I | Y | H | S | V | S | H | A | R | P | R | W | I | W | F | C | L | L | L | L | A | A | G | V |
| 7201 | GGGACGATT | TATCAGAGC | TGTCTCATC | CCGCGCCGC | TGGATCTGGT | TTTGCCTACT | CTGTCTTGT | GCAGGGGTAG | CCCTCTGTAA | ATAGTGTGCG | ACAGAGTAGG | GGCCGGGGG | ACCTAGACCA | AAACGGGATGA | GGACGAAGGA | CGTCCCATC | | | | | | | | | | |

| | | | | | | | | |
|------|------------|-----------|-----------|-----------|-----------|-----------|------------|-----------|
| +2 | G | I | Y | L | L | P | N | R |
| 7281 | CGATCTACCT | CCTCCCAAC | CGATGAAGT | TGGGGTAAC | ACTCGGCGT | AAAAAATA | AAAAATCTAG | AAAGCGCGC |
| | CGTAGATGA | GGAGGGGTG | GCTACTTCA | ACCCATTG | TGAGCGCGG | TTTTTTTTT | TTTTTAGATC | TTTCGGCGC |

| | | |
|--|-------|-------|
| | BANHI | MluI |
| | ----- | ----- |

| | | | | | | | | |
|------|------------|------------|------------|------------|------------|------------|------------|------------|
| 7361 | CAAGATATCA | AGGATCCACT | ACCGGTTAGA | GCTCGCTGAT | CAGCTCGAC | TGTGCCTTCT | AGTGGCCAGC | CATCTGTTGT |
| | GTCTTATAGT | TCCTAGTGA | TGGCAATCT | CGAGCGACTA | GTCGGAGCTG | ACAGGGAAGA | TCAACGGTCC | GTAGACAACA |

| | | | | | | | | |
|------|------------|------------|------------|------------|-----------|-------------|------------|------------|
| 7441 | TTGGCCCTCC | CCCGTGCCTT | CCTTGACCT | GGAAGTGCC | ACTCCACTG | TCCTTTCCCTA | ATAAATGAG | GAATTTGAT |
| | AACGGGGAG | GGGCACGGA | GGAACCTGGA | CCTTCCACGG | TGAGGTGAC | AGGAAGGAT | TATTTTACTC | CTTTAACCTA |

FIG. 3-Page 13

pCMV-NS3S

| | |
|------|--|
| 7521 | GGCATTGTCT GAGTAGTCT CATTCTATT TGGGGGGTGG GTGTGGGAG GACAGCAGG GGGAGGATTG GGAAGACAAT CGCTAACAGA CTCAATCACA GTAAATAG ACCCCGACG CGACCCGCTC CTCTGGTTCC CCGCTCCTAAC CTTCTGTGTTA |
| 7601 | AGCAGGCATG CTGGGAGCT CTTCGGCTTC CTGGCTCACT GATCGGCTGC GTCGGTGGT TCGGCTGGG CGAGGGGTAT TCGTCCGTAC GACCCCTCGA GAAGGGGAAG GAGCGAGTGA CTGAGCGAG CGAGCCAGCA AGCCGACGCC GCTCGCCATA |
| 7681 | CAGTCACTC AAAGGGGGTA ATACGGTAT CCACAGAATC AGGGGATAAC GGAGAAAGA ACATCTAGC AAAGGCCAG CTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCTCTATT GTCCTTTCT TGTACACTCG TTTTCGGGTC |
| 7761 | CAAAAGGCA GGAACGGTHA AAAGGGGGG TTCTGGGGT TTTTCATAG GCTCGGCCCC CCTGACGAGC ATCACAAAA GTTTTCCGCT CTTGGCATT TTTCCGGGCG AACGACGCA AAAAGGTATC CGAGGCGGG GGACTGGCTG TAGTGTITTT |
| 7841 | TCGACCTCA ACTCAGACT GCGGAAACC GACAGACTA TAAAGATACC AGCGTTTTC CGCTGGAAGC TCGCTCGTGC AGCTCGAGT TCACTCTCA CCGCTTTGGG CTCTCTCAT ATTCTATGG TCCGCAAGG GGGACCTTCG AGGAGACAG |
| 7921 | GCTCTCTGT TCGGACCTG CCGCTTACCG GATACCTGT CCGCTTCTC CTTTGGGAA GCGTGGCGCT TTCTCAATGC CGAGAGACA AGCTGGGAC GGGGAATGG CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGATTACG |
| 8001 | TCAGGCTGA GGTATCTAG TTGGGTGAG GTGTTTGGT CCAAGTGGG CTGTGTGAC GAAACCCCGG TTCAGCCGGA AGTGGACAT CCATAGAGT AGCCACATC CAGCAAGCGA GGTTCGAGCC GACACACGTG CTTGGGGGGC AAGTCGGGCT |
| 8081 | CCGTGGGCG TTATCCGGTA ACTATGCTT TGAGTCCAAC CCGGTAGAC AGACTTATC GCGACTGGCA GGAGCCACTG GGCGACCGG AATAGGCCAT TGATAGCAGA ACTCAGTTG GGCAATCTG TGTGAATAG CCGTGACCGT CGTGGGTGAC |
| 8161 | GTACAGGAT TAGCAGGCG AGTATGTAG CCGGTGCTAC AGACTTCTT AGTGTGGC CTAACCTACG CTACACTAGA CATTTGCTTA ATGCTCTCG TCCATACATC CGCCAGGATG TCTCAGAAC TTCAACCCG GATTGATGCC GATGTGATCT |

FIG. 3-Page 14

pCMV-NS35

| | |
|------|---|
| 8241 | AGGACAGTAT TTGGTATCTG CGCTCTGGTG AAGCCAGTTA CCTTGGGAAA AAGAGTTGGT AGCTTTGAT CGGCGAAACA TCCTGTCATA AACCATAGAC GCGAGACGAC TTGGTCTCAAT GGAAGCCCTT TTCTCAOCCA TCGAGAAGCTA GGCCTTTTGT |
| 8321 | AACACACGGT GGTAGGGGTG GTTTTTTCT TTGCAAGCAG CAGATTAGCG GGAGAAAAA AGGATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGGCCA CAAAAAACA AAGTTTGGTC GTCTAATGCG CGCTTTTTT TCTTAGAGTT CTCTAGGAA |
| 8401 | TGATCTTTTC TACGGGTCT GAGCTCACT GGAACGAAAA CTCAGCTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ACTAGAAAAG ATGCCCCAGA CTGGAGTCA CCTTGGCTTT GAGTGCATTT CCTAAAAACC AGTACTCTAA TAGTTTTTTC |
| 8481 | ATCTTCACTT AGATCCTTTT AAATTAAAA TGAAGTTTAA ATCAATCTA AGTATATAT GAGTAAACTT GGTCTGACAG TAGAAGTGGA TCTAGGAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCAATATAA CTCAATTGAA CCAGACTGTC |
| 8561 | TTACCAATGC TTAATCAGTG AGGACCTAT CTCAGGCATC TGCTCTATTC GTTCATCCAT AGTTGGCTGA CTCGCCGTGC AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGCAGC |
| 8641 | TGTAGATAAC TACGATACG GAGGCTTAC CATCTGCCCG CAGTGTGCA ATGATACGGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCGAATG GTAGCCGGG GTACGACGT TACTATGCGG CTCGGGTGC GAGTGGCGGA |
| 8721 | CCAGATTAT CAGCAATAA CCAGCCAGCG GGAAGGGCG AGGCAGAAAG TGCTCCTGCA ACTTTATCCG CTTCATCCA GGTCTAAATA GTCTTATTT GTCTGGTGG CCTTCCCGCG TCGGCTCTC ACCGAGGAGT TGAATATGGC GGAGGTAGGT |
| 8801 | GTCTATTAAAT TGTTCGGGG AAGTAGAGT AAGTAGTTG CCAAGTTAATA GTTTCGGCAA CGTTGTTGCC ATTGCTACAG CAGATAANTA ACAACGGCC TTGCTATCTA TTCAATCAAG GGTCAATTAT CAAACGGCTT GCAACAACCG TAACGATGTC |
| 8881 | GCATCTGGT GTACAGGTG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACCAT CAAGGGAGT TACATCATCC CTTAGCACCA CAGTGGGAGC AGCAAAACCAT ACCGAAAGTAA GGTGAGGCGA AGGTTTGCTA GTTTCGGCTCA ATGTAAGT |

FIG. 3-Page 15

pCMV-NS35

8961 CCCATGTTGT GCAAAAAGC GGTAGTCTC TTGGTCTCT CGATCGTGT CAGAGTAG TTGGCGGAG TGTATCACT
GGGTACAACA CTTTTTTCG CCAATCGAG AAGCAGGAG GGTAGCAACA GTTTCATTC AACGGCGTC ACAATAGTGA

9041 CATGCTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCGCTAA GATCGTTTTT TGTGACTGCT GACTACTCAA
GTACCAATAC CGTCGTGAG TATTAAGAGA ATGACAGTAC GGTAGCAATT CTACGAAAG ACATGACCA CTCATGAGTT

9121 CCAAGTCATT CTGAGAATAG TGTATGGGC GACGAGTTG CTCTTCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT
GGTTCAGTAA GACTCTTATC ACATACCGCG CTGGCTCAAC GAGAAAGGC CGAGTTATG CCCTATTATG CGCGGTGTA

9201 AGCAGAACTT TAAAAGTGCT CATCATTTGA AAAGCTTCT CGGGGGGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC
TGTCTTTGAA ATTTCAGGA GTAGTAACCT TTTGCAAGAA GCGCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTGATG TAACCCACTC GTGACCCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTCTGGG TGACAAAAA
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAT GAAAGTGTG GCAAGAGCCC ACTCGTTTTT

9361 CAGGAAGCCA AAATGCCGCA AAAAAGGGA TAAGGGCGAC ACGGAATGT TGAATACTCA TACTCTTCT TTTTCAATAT
GTCTTCCGT TTTACGGCGT TTTTCCCTT ATTCCGCGTG TGGCTTTACA ACTTATCAGT ATGAGAAAGA AAAAGTTATA

9441 TATTGAAGCA TTTATCAGG TTTTGTCTC ATGAGCGAT ACATATTTCA ATGTATTAG AAAATAAAC AAATAGGGT
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAACT TACATAATC TTTTATTG TTTATCCCCA

9521 TCCGGCACA TTTCCCGAA AAGTGCACG TGAGCTTAA GAAACATTA TTATCATGAC ATTAACCTAT AAAATAGGC
AGGCCGTGT AAAGGGCTT TTAACGGTGG ACTGCAGATT CTTTGTGTAAT AATGACTG TAATTGGATA TTTTATCCG

9601 GTATCAGGAG GCCCTTTCGT C
CATATGCTC CGGAAAGCA G

FIG. 3-Page 18

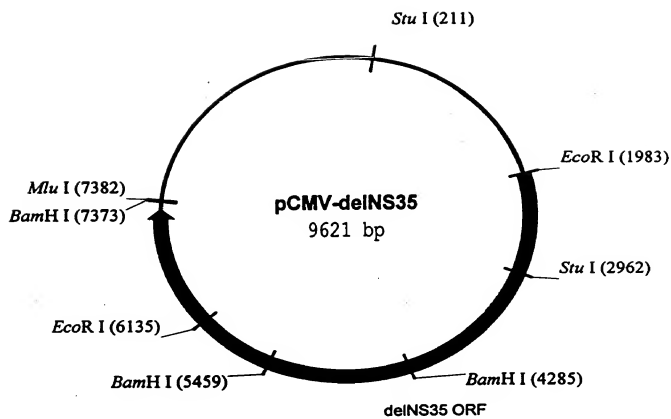


FIG. 4

pCMV-deINS35

| | |
|-----|---|
| 1 | TCGGGGTTT CGGTGATGAC GGTGAAMACC TCTGACACAT GCAGCTCCCG GAGAGGGTCA CAGCTTGTCT GTAAGCGGAT AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTCTGTA CGTCGAGGCG CTCGCCAGT GTCGAACAGA CATTCGGCTA |
| 81 | CCCGGAGCA GACAAGCCG TCAGCGCGG TCAGCGGGTG TTGGCGGGTG TCGGGGTGG CTAACTATG CGGCATCAGA CGGCGCTCGT CTGTGGGCG AGTCGCGGCG AGTCGCGCAC AACGCGCAC AGCGCCGACC GAAATTCATAC GCGGTAGTCT |
| | ----- StuI |
| 161 | GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTCA AAAGCTAGG CTCCAAAAA AGCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACT CGAAAAAGGT TTTGGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGAC |
| 241 | AATAGCTCAG AGCGCGAGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG ATTGGGGGA TTATCGAGTC TCGGGCTCG CGGAGCGCG AGACGTATTT ATTTTTTTA ATCAGTCGT ACCCGGCTC TTACCGGCT |
| 321 | ACTGGGGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACCTTGATC TATATCATAA TATGTACATT TATATTGGCT TGACCGGCC CTCCTTAAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACGGA |
| 401 | CATGTCGAAT ATGACGGCA TGTTGACATT GATTATTGAC TACTTATTAA TAGTAATCAA TTACGGGTC ATTAGTCAT GTACAGETTA TACTGGCGGT ACAACTGTAA CTAATNACTG ATCAATAATT ATCATIAGTI AATGCCGAG TAATCAAGTA |
| 481 | AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCGGCC TGCGTGACCG CGCAACGACC CGCGGCCATT TCGGGTATAT ACCTCAAGC GCAATGTATT GAATGCCATT TACCGGGCGG ACGACTGCG GGGTTCTGG GGGCGGGTAA |
| 561 | GAGTCAATA ATGACGTATG TTCCCATAG AAGGCCAATA GGCACCTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCAATC AAGGGATCA TTGGGGTTAT CCGTGAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA |

FIG. 5-Page 1

CMV-delNS35

641 AAATGCCCA CTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCGGCC CCTATTGAC TCAATGACGG TAAATGGCC
TTTAGCGGT GAACCGTCAI GTAGTICACA TAGTATACGG TTCAGCGGG GGAATCTGC AGTACTGCC ATTACCGGG

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTACTTGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
CGGACGCTAA TAGCGGTCAI CTACTGGAAT GCGCTCAAG GATGAACGGT CATGTAGATG CATATCACT ACGGATAATC

801 CATGTGTATG CGGTTTGGC AGTACACCAA TGGGGCTGGA TAGCGTTTG ACTCACGGG ATTTCCAAAT TTCCACCCCA
GTACCACTAC GCCAAACCG TCATGTGGT ACCCGCACT ATCGCAAAAC TGAATGCCCC TAAAGTTCA GAGTGGGT

881 TTGACGTCAA TGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTGTA ATACCCCGC CCGTTGACG
AACTGCAGTT ACCCTCAAC AAAACCGTGG TTTTATTTGC CTGAAAGGT TTACAGCAT TATTGGGGCG GGGCACTGC

961 CAAATGGGG GTAGGCGTGT ACGTGGAG GTCTATATA CGAGAGCTCG TTTAGTGAAC CCTGAGATCG CCTGGAGACG
GTTTACCCGC CATCCGACA TGCCACCTC CAGATATAT CGTCTGAGC AAATUACTTG CGAGTCTAGC GGACTCTGC

1041 CCATCCACG TGTTTGGACC TCATAGAAG ACACGGGAC CGATCCAGCC TCGGGGGCG GGAAGGGTGC ATTGGAACGC
GCTAGGTGG ACAAACTGG AGTATCTTC TGTGGCCCTG CTTAGGTGG AGCGCGCGC CCTTGGCAGC TAACTTTGG

1121 GGATTCGGG TGCAGAGCT GAGTAACTA CGGCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGGC ACGTTCTCA CTGCATTCAI GCGGATATC TGAGATATCC CTGTGGGGA ACCGAGAATA CCGTATAT

1201 CTGTTTTGG CTTGGGGCT ATACACCCCC GTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTGA
GACAAAAAC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT

1281 TTGACCATTA TTGACACTC CCTATTGGT GACGATACTT TGCATTACTA ATCCATACA TGGCTCTTTG CGACAACAT
AACTGCTAAT AACTGTGAG GGGATAACCA CTGCTATGAA AGTAACTAT TAGGTATTGT ACCGAGAAC GGTGTTCATA

FIG. 5-Page 2

pCMV-deINS35

| | |
|------|---|
| 1361 | CTCTATTGGC TATATGCCAA TACTCTGTCC TTGAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTATTAT GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA |
| 1441 | TATTTCACAA TTCACATATA CACAACGGC GTCCCCCGTG CCGCGAGTTT TTAATTAACA TAGCGTGGGA TCTCCGACAT ATAAATGTTT AAGTGTATAT GTTGTTCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGACCCCT AGAGGTGTA |
| 1521 | CTCGGGTACG TGTTCGGGAC ATGGGCTCTT CTCCGCTAGC GCGCGAGCTT CCACATCGGA GCCTTGCTCC CATCGTCCCA CAGGCCATGC ACAAGGCGTG TACCCGAGAA GAGGCCATCG CCGCTTCGAA GGTGTAGGCT CCGGACCAGG GTAGGCAGGT |
| 1601 | CCGGCTCATG CTCGCTCGGC AGCTCCTTGC TCCTAACACT GGAGGCGAGA CTTAGGCACA GCACAATGCC CAGCACCCACC CGCGGAGTAC CAGGAGCCCG TCAGAGAACG AGCATTTCTA CTTCGCGTCT GAATCCGTGT CGTGTACGG GTGCTGTGG |
| 1681 | AGTGTGCGGC ACAAGGCGGT GCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACTT GGACGCGAGAT TCACACGGCG TGTTCGGGCA CCGGCATCCC ATACACAGAC TTTTACTCGA GCCTCTAAC CCAGCGTGA CTTGCGTCTA |
| 1761 | GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCGAGT GAGTCTTTGT ATTCTGATAA GACTCAGAGG TAACTCGCGT CCTTCTGAAT TCCGTGCGCG TCTTCTCTA CGTCCGTCGA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA |
| 1841 | TGCGGTGCTG TTAACGCTGG AGGCGAGTGT AGTCTGAGCA GTACTGCTTG CTGCGCGCGG CGCCACCAGA CATAATAGCT ACGCCACGAC AATTGCCACC TCCGCTCACA TCAGACTGCT CATGAGCAAC GACGGCGGCG CGGCTGGTCT GTATTATCGA |
| +2 | |
| 1971 | GACAGACTAA CAGACTCTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT GTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACCTA |

EcoRI

M A A

FIG. 5-Page 3

pCMV-deINS35

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGGTATAAG GTGTAGTAC TCAACCCCTC TGTGTGCGA ACATGGGCT TTGGTGCTTA CATGTCCAAG
TACGTGAGT CCCGATATTC CAGCATCATG AGTTGGGAG ACAACGAGT TGTGACCGA AACACGAAT GTACAGTTC

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGA TCGATCCTAA CATCAGACC GGGGTGAGAA CAATTACAC TGCAGCCCC ATCAGTACT CCACCTACGG
CGAGTACCT AGCTAGGATT GTAGCTCTG CCCCACTTT GTTAATGGTG ACGTGGGG TAGTGCATGA GTGGAGTCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCAGCGGG GTGCTCGG GGGCGTTAT GACATAATAA TTTGTGAGA GTCCCACTCC ACGATGCCA
GTTCAAGGAA CGGCTGCCG CCAGGAGCC CCGCGAATA CTGTATTAT AAACACTGCT CAGGTGAGG TGCTACGCT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCATCTT GGCATTGGC ACTGTCTTG ACCAAGCAGA GACTGGGG GCGAGCTGG TTGTGCTGC CACCGCCACC
GTAGTAGAA CCGCTAACGG TGACAGGAAC TGGTTGTCT CTGACGCC CGCTCTGACC AACACGAGC GTGGCGGTG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCGGGCT CGTCACTGT GCCCATCC AACATCGAG AGTGTCTT GTCCACCACC GGAGATCC CTTTTAGCG
GGAGGCCCGA GCAGTGACA GGGGTAGG TTGTAGTCC TCCAAGAGA CAGGTGGTG CTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGCTATC CCCCTCAAG TAATCAAGG GGGGACAT CTGATCTT GTCAATCAA GAAGAAGTC GAGCAACTCG
GTTCCGATG GGGGAGTTC ATTAGTTCC CCCCTCTGA GAGTAGAGA CAGTAAGTT TTTCTTACG CTGCTTGAG

+2 A A K L V A L G I N A V A Y Y R C L D V S V I P T S G
2481 CCGCAAGCT GTGCGATTG GGCATCAATG CCGTGGCTA CTACCGGT CTTCAGTGT CCGTATCC GACGAGCGC
GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACGGAT GATGGGGCA GAACCTGACA GGCAGTAGG CTGCTGGCG

FIG. 5-Page 4

pCMV-delINS35

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
 2561 GATGTTGCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATAGCGGCGA CTTGCACTCG GTGATAGACT GCAATACGCT
 CTACACACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCCGT GAAGCTGAGC CACTATCTGA CGTTATGAC

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 TGTACCCAG CACTGCGATT TCAGCCCTGA CCCTAGCTTC ACCATTGAGA CAATACGCT CCCCAAGAT GTGTCTCCG
 ACATGGGGTC TGTCAGCTAA AGTCGGAAT GGGATGGAAG TGTTAACTCT GTTAGTGGA GGGGGTTCTA CGACAGAGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCATCAACG TCGGGGAGG ACTGGCAGG GGAAGCCAG CATCTACAGA TTTGTGGAC CGGGGAGCG CGCTCCGCG
 CGTAGGTTGC AGCCCCGCTC TGACCGTCC CTTCCGTC GTAGATGCT AAACACCTG CCCCCCTGC GGGGAGGCG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTGACT GTCGCTCT CTGTGATCG TATGAGCAG GCTCTGTTG GTATGAGCTC ACGCCGCGCG AGACTACAGT
 TACAAGCTGA GCAGGCGAGA GACACTCAG ATACTGGTC CGACACGAAC CATACTCAG TGGGGGCGCG TCTGATGTA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 2881 TAGGCTACGA CGGTACATGA ACACCCCGG GCTTCCGCTG TGGCAGGAC ATTTGAAT TTGGGAGGCG GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGCCC CGAAGGGCAC ACGTCTCTG TAGAATTAA AACCTCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q

 StuI
 2961 GCCTCACTA TATAGTGGC CACTTTCTAT CCCAGACAAA CGAGGTGGG GAGAACCTTC CTTACTGCT AGGCTACCAA
 CGGAGTGAAT ATATCTACGG GTGAAGATA GGGTCTGTTT CGTCTCAACC CTCTTGGAA GAATGGACCA TCGCATGGT

FIG. 5-Page 5

pCMV-delNS35

| | |
|-------|---|
| +2 | A T V C A R A Q A P P S W D Q M W K C L I R L K P T |
| 3041 | GCCACCTGT CCGTAGGC TCAAGCCCT CCCCATCGT GGGACAGAT GTGGAAGTGT TTGATTCCGC TCAAGCCAC |
| | CGGTGGACA CGGATCCG AGTTCGGGA GGGGTAGA CCTGTGCTA CACCTTACA AACTAAGCG AGTTCGGGTG |
| <hr/> | |
| +2 | L H G P T P L L Y R L G A V Q N E I T L T H P V T K |
| 3121 | CTCCATGGG CCAACACC TCCTATACG ACTGGCGT GTTCAGATG AATCACCT GAGCAGCCA GTCAACCAAT |
| | CGAGTACC GGTGTGGG ACGATATGTC TCACCCGCA CAAGCTTAC TTAGTGGG CTGCTGGGT CAGTGTTTA |
| <hr/> | |
| +2 | Y I M T C M S A D L E V V T S T W V L V G G V L A A L |
| 3201 | ACATCATGAC ATGCATCG CCCACCTGG AGTCTGTCAC GAGCAGCTGG GTGCTGTTG GCGCGCTCT GGTGCTTTG |
| | TGTAGTACTG TACGTACG CCGCTGGACC TCCAGCAGTG CTCTGGACC CACGAGCAAC CGCCGCGGA CCGACGAAAC |
| <hr/> | |
| +2 | A A Y C L S T G C V V I V G R V V L S G K P A I I P D |
| 3281 | CGCGGTATT GCCTCAAC AGGCTCGTG CTCATAGTG GAGGGTCTT CTTCGCGG AAGCGGCAA TCATAGCTGA |
| | CGCGCATAA CGGACGTTG TCCAGCCAC CAGTATACC CGTCCAGCA GAACGCGCC TTGCGCGCTT AGTATGGACT |
| <hr/> | |
| +2 | R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M |
| 3361 | CAGGCAAGTC CTCTACGAG AGTTCATGA GATGGAAGG TCCTCTCAG ACTTACGCTA CATCGAGCAA GGGATGATG |
| | GTCCTTACG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTG TGAATGGCAT GTAGCTCGTT CCCTACTAGG |
| <hr/> | |
| +2 | L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V |
| 3441 | TCCGCGACA GTTCAAGCAG AAGCCCTCG GCCTCTGCA GACGCGCTCG CGTCAGCAG AGGTATCG CCCTGTGTG |
| | AGCGCTCT CAAGTTGCTC TTCGCGAGC CGGAGGACGT CTGCGCGAGG GCATCCCTC TCCAAATAGG GGGACGACAG |
| <hr/> | |
| +2 | Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G |
| 3521 | CAGACCAAT CGCAAACT CGAGACTTC TGGCGAAGC ATATGTGAA CTTCATCAGT GGGATACAAT ACTTGGCGG |
| | GCTCGTTGA CGCTTTTGA GCTCTGGAAG ACCGCTTCG TATACACTT GAAGTAGTCA CCCTATGTTA TGAACGCGCC |

FIG. 5-page 6

pCMV-delNS35

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
 3601 CTGCTCAAG CTGCTGCTA ACCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAAGCCCA CTAACACCTA
 GAAACGTTGC GACGAGCCAT TGGGGGGGTA ACGAAGTAAC TACCGAAAT GTCGACGACA GTGCTCGGT GATTGCTGAT

 +2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
 3681 GCCAAACCT CCTCTCAAC ATATTGGGG GTGGGTGGC TGCAGCTC GCGGCCCGC GTGCGGCTAC TGCTTTCTG
 CGCTTTGGGA GGAGAAATTG TATAACCCG CCACCCACCG ACGGCTCAG CCGGGGGGGC CACGGCGATG ACGAAACAC

 +2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
 3761 GCGCTGCT TAGTGGCG CGCATCGC AGTGTGGAC TGGGAAGT CCTCATAGAC ATCCTTGCAG GGTATGGCG
 CCGGACCGA ATCGACCGG CGGTAGCG TCAGAACCTG ACCCTTCCA GGAATATCTG TAGGAACGTC CCATACCGC

 +2 G V A G A L V A F K I M S G E V P S T E D L V N L L
 3841 GGGGTGGG GAGCTCTTG TGGATTCAA GATCATGAGC GTTGAGTCC CCTGCACGA GGACCTGTC AATCTACTGC
 CCGCACCGC CCTCGAGAC ACCGTAAGT CTAGTACTG CCACTCCAGG GGAGTGCTT CCTGGACGAG TTAGATGAGC

 +2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
 3921 CCGCATCT CTGCGCGGA GCGCTGTAG TCGGCTGTG CTGTGCAGCA ATACTGCGC GGCAGTTGG CCGGGCGAG
 GGGCTAGGA GAGCGGCT CCGGAGCATC AGCGGACCA GACACTGCT TATGACGCG CCGTCCAACC GGGCGCGCT

 +2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
 4001 GGGCAGTGC AGTGGATGAA CCGCTGATA GCCTTGGCT CCGGGGAA CCATGTTCC CCGAGGCACT AGTCCCGA
 CCGCTCAG TCACCTACTT GCGCGACTAT CGGAGCGGA GGGCGGCTT GGTACAAAG GGTGCTGA TGCACGGCT

 +2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
 4081 GAGCATGA GCTGCCCGG TCATCCCAT ACTCAGCAGC CTCACTGTAA CCGACTCT GAGGGGACTG CACCAAGTGA
 CTGCTACT CTGACGGGCG AGTGAAGGTA TGAAGTCTG GATGACATT GGTGAGGA CTCGCTGAC GTGCTCACT

FIG. 5-Page 7

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
 +161 TAACTCGGA GTGTACCACT CCATGCTCCG GTTCTCTGCT AAGGACATC TGGCATCTGA TATCGAGGT GTTGAAGCAC
 ATTGCGCCT CACATGCTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGGTG

4241 TTTAAGACCT GGCTAAAGC TAAGTCA TG CACAGTGC CTTGGATCC CTTTGTCTCC TGCACGCGG GGTATAAGG

+2 V W R G D G I M H T R C H T C T G C G A E I T G H V K N G T
 4321 GGCTCGCGA TCATGCACG TCCTGCCAC TGTCGACGTG AGATCACTGG ACATCTCAAA AAGCGACGA
 CGATACCGCT CCGCTCGCGT CAGTCTGTG AGGCACGGTG ACACCTCGAC TCCTAGTACC TGTACAGTT TGTGCCCTGCT

+2 M R I V G P R T C R N W S G T F P I N A Y T T G C P C
4401 TCAGATCGT ACCTGCTAGG ACCTGCTAGG ACATGCGG TGGACCTTC CCCATTAATG CCTACACAG GGGCCCGCTG
4402 CCGCTCTAGA CGCAGGATCG TGGCAGCTCT TGTACAGGTC ACCCTGGAAG GGGTAAATG GCGATGTGTG CCGGGCGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
 4481 ACCCGCTTC CTGCGCGAA CTACAGTTC GGTGTGCG AGAGAAATAC GTGAGAGTAA GCGAGTGGG
 TGGGCGGAG GACGGGGT CATGTGAAG CGCGATACG CCACAGACG TCTCTTATG CCGTCTATT CCGTCCACCC

4561
+2
D F H Y V T G G M T T D N L K C P C Q V P S P E F F T C
GCTATCCAC TACGTACGG GTATCTATC TGACATCT AAATGCCGT GCCAGTCCT ATCGCCGAA TTTTTCAC
CGTGAAGTG ACTGACTGC CATACTATG CACTGTAGAA TTATCGGCA CGCTCCAGC TAGGGGGT AAAAAGTGC

+2 E L D G V T R L H R F A P F C K P L L R E E V S F R V G
4641 AATLGAAGG GGTGGCCTA CATAGTTTG CGCCCCCTG CTGCGGGAG AGGTATCAT CAGAGTAGGA
TTAACTGCG CCACGCGAT GTATCCAAAC GCGGGGGGAC GTTTCGGGAC CAGCGCTCG TCCATAGTAA GTCTCATCTCT

pCMV-delNS35

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCACGAAT ACCGCTAGG CTGGCAATTA CTTGCGAGC GCGAACCGGA CTGGCCGCTG TTGAGTGCA TGCTCACTGA
GAGTGCTTA TGGGCATCC GAGGTTAAT GGAACGCTCG GCGTTGGCTT GACCGGCAC AACTGCAGT AGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A
4801 TCCTCCCAT ATACAGCAG AGCGCGCG GCGAGTTG CCGAGGGAT CACCCGCTC TGTGGCCAGC TCCTCGGCTA
AGGAGGCTA TATTGCTGC TCCGCGCGCC CGCTTCCAA CCGTCCCTA GTGGGGGAG ACACGGTGC AGGAGCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N
4881 GCCAGTATC CGCTCCATCT CTCAAGCAA CTTGCACGC TAACATGAC TCCCTGATG CTGAGTCTAT AGAGCCCAAC
CGCTCGATAC GCGAGTAGA GAGTCCGTT GAACGTGGCG ATTGGTACTG AGGGACTAC GACTCGAGTA TCTCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGA GCGAGGAT GCGCGCAAC GATCTCCGTA CCGCAGAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCTCTGA
GAGGATACCT CGCTCTCTA CCGCGCGTTG TAGTGTCCC AACTGAGTCT TTGTTTAC CACTAAGACC TGAGGAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCCTTTGT GCGAGGAGG ACAGCGGGA GATCTCCGTA CCGCAGAAA TCCTGCGAA GTCTGGAGA TTCGCCGAG
AGCGAACAC CGCTCTCTC TGCTCGGCT CTAGAGGAT GGGCGTCTT AGAGCGCTT CAGAGCTCTT AAGCGGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCTCGGCT TTGGCGGG CCGAGTATA ACCCGCGCT ACTGGAGCTT GCGAAAGC CCGACTAGA ACCACTTG
GGAGCGGGA AACCCGCG GCGCTGATAT TGGGGGGGA TCACTCTGC ACTTTTTC GGTGATGCT TGTGGAGAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E
5201 GTCCATGGCT GCGCGTTCC ACTCTCAAG TCCTCTCTG TGCTCCGCG TCGGAAGAAG CCGAGGTTG TCCTCACTGA
CAGGTACCA CCGCGGAAG TGAAGTTTC AGGGAGGAC ACGGAGCGG ACCTTCTTC GCTGCCACC AGAGTGACT

FIG. 5-Page 9

pCMV-delNS35

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D
5281 ATCAACCTA TCTACTGCT TGGCGAGT CGGACACGA AGTTTGGCA GTCTCTCAAC TTCGGCAAT ACGGCGGACA
TAGTTGGGAT AGATGACGA ACCGGTCTGA CGGTGTCT TCGAAACCGT CGAGGAGTTG AAGGCGGTAA TGCGCGGTCT

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATAGACAAC ATCTCTGAG CCGCGCCCTT CTGGCTGCC CCGGACTCC GACCTGAGT CCTATTCTC CATGCCGCC
TATGCTTTG TAGGAGACTC GGGGGGGGAA GACGACGG GGGGTGAGG CTGCACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V
BamHI

5441 CTGAGGGGG AGCCTGGCA TCCGATCTT AGCGAGGGT CATGGTCAAC GTTCACTAGT GAGGCCAACG CGGAGGATGT
GACCTCCCC TCGGACCCCT AGGCTAGAA TCGTGCCTA GTACGAGTTG CCACTCATCA CTCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGTGGTG TCAATGTCTT ACTCTTGAC AGCGCACTC GTCAACCCCT GCGCGCGGA AGAACAGAAA CTGCCCATCA
GCACACGAGC AGTTACAGAA TGAGAACCTG TCGCGGTGAG CAGTGGGGCA CGCGGGGCT TCTTGTCTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCACTAAG CAATCTGTG CTAGTCAAC ACAATTTGT GTATTCCAC ACCTCACGCA GTCTTTGCCA AAGCGAGAAG
TAGCTGATTC GTTGAGCAAC GATGAGTGG TGTTAAACCA CATAAGTGG TGGAGTGCT CAGGAACGCT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
5681 AAAGTCACT TTGACAGCT CCAAGTTCTG GACAGCACT ACACAGGCT ACTCAAGGAG GTTAAAGAG CGCGCTCAAA
TTTCAGTGTA AACTGTCTGA GTTTCAGAGC CTGTGCGTAA TGGTCTGCA TGAGTTCTC CAATTTCTC CGCGGAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AGTGAAGCT AACTTGGAT CCGTAGAGA AGCTTCAGC CTGACGCCCC CACACTCAGC CAAATCCAG TTGTGTATG
TCACTTCCGA TTGAACGATA GGCACTCTCT TCGAAGCTGG GACTGGGGG GTGTGAGTGG GTTAGGTTT AAACCAATAC

FIG. 5-Page 10

pCMV-deINS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGCAAGA GTCCCTTGG CATCCAGAA AGCGCTAAC CCACATCAAC TCCTGTGGA AACCTTCT GGAGACAAT
CCCGTTTCT CGAGGCAAG GTAGGTCTT TCGGCAITG GGTGTAGTG AGGACAACCT TTTGGAAGA CTTTGTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
5921 GTACACAA TAGACACTAC CATCATGGCT AAGACGAGG TTTTCTGGCT TCAGCTGAG AAGGGGGTC GTAAGCCAGC
CATTTGGTT AICGTGTGAT GTAGTACCA TTCTTGCTC AAAAGACGA AGTCGGACTC TTTCCCCAG CATTTGGTGT

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCCTCTATC GTGTTCCCG ATCTGGGCT GCGGTGTC GAAAAGATGG CTTTGTAGA CTTGTTTACA AAGCTCCCT
ACGAGTAG CACAAGGGC TAGACCGCA GGGCACAG CTTTCTTACC GAAACATGCT GCACCAATGT TTTGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

6081 TGGCGGTGAT GGAAGTCC TAGGATTCC AATACTCAC AGGACAGGG GTTGAATTCC TCCTGCAAGC GTGGAAGTCC
ACCGGCACTA CCTTCGAG ATGCTAAGG TTATGAGTG TCCTGTGCGC CAACCTAAGG AGCAGTTCC CAGCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAACCC CAATGTGGT CTGTAATG ACCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA
TTCTTTTGGG GTTACACCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCCTGTAGG CATGCTCTCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGCAATCTAC CAATGTGTG ACCTGACCC CCAAGCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTGGG
CCGTTAGTG GTTACAAAC TGGAGTGGG GGTTCGGCG CACCGGTAGT TCAGGGAAGT GCTCTCCGAA ATACAAACCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTCTAC CAATTCAAG GGGAGAAT CGGCTATCG CAGGTCCGC GCGAGCGCG TACTGACAC TACTGTGTG
CGGAGAATG GTTAAGTTC CCCTCTTGA CGCGGATAG GTCCAGGCG CGGTGCGCG ATGACTGTTG ATCGACACA

pCMV-delINS35

+2 N T I T C Y I K A R A A C R A A G L Q D C T M L V C G
6401. AACACCTCA CTTCGTACAT GAGCGCGG GCAGCTGTC GTCCAGCAG TCACCATGC TCCTCTGTG
TTCTGGGAGT GAACGATGA TTTCGGGGC GTCGGAGC GTGGGGTCC GAGGTCTG AGTGGTAGC AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481. CGACGACTA GTGCTTATCT GTGAAGCGC GGGGTCCAG CAGCTGACG AGCTTCACG GAGCTATGA
GCTGCTGAAT CAGCAATAGA CACTTTGCG CCCCAGGTC TCCTGGGCC GTCGGACTC TCGGAAGTC GTCCGAIAC

+2 T R Y S A P G D P P Q P E Y D L E L I T S C S S N V
6561. CCAGGTACTC CGCCCCCTT GGGACCCGC CACAACGAG ATACCACTTG GAGCTCATAA CATCATGTC CTCGAGGTG
GTTCCATGAG CGGGGGGGGA CCGCTGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTAT GTAGTACGAG GAGTTGAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641. TCACTGGCC ACAGCGGCG GTCTACTACC TCACCCGTGA CCGTACACCC CCGCTCGGA GAGCTGGTG
AGTCAGCGG TGCTGGCGG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGAGCGCT CTCGACCGAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721. GGAGCAGA AGACACTC CAGTCATTC CTGGCTAGC AACATACTA TGTTCGCCC CACACTGCG GCGAGGATGA
CCTCTGCT TCTGTCGAG GTCACTAAG GACCGATCG TTGATTAAGT ACAAACGGG GTGTGACACC CGCTCTTACT

+2 I L M T H F S V L I A R D Q L E Q A L D C E I Y G A
6801. TATGATAGC CAATTCCTT AGCTGCTTA TACCGAGGA CAGGCGCTCG ATTGCGAGT CTACGAGGCC
ATGACTACTG GGTAAAGAA ATCGGTCCCT GTTCGAACTI GTCCGGGAGC TAACGCTCTA GAIGCCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881. TGCTACTCA TAGAACCTT GGTACTACT CCAATCATT AACAGTCA TGGCTCAGC GCAATTCAC TCCAGCTTA
ACGACTAGT ATCTTGTTGTA CTAGATGGA GTTATGAAG TTTCTGAGT ACCGAGTCG CGTAAAGTG AGGTCTAA

FIG. 5-Page 12

PCMV-delNS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
6961 TTCTCCAGT CAATACATA GGTGGCCGC ATGCCTCAGA AAATTTGGG TACGGCCTT CGACGCTTGG AGACACCGGG
GAGAGGTCCA CTTTAGTTAT CCCACCGGG TACGGAGTCT TTGGAACCCC ATGGCGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
7041 CCCGAGCCT CCGCGCTAGG CTTCCTGCCA GAGAGGCG GGTTCGCCAT ACCTTCAA CTGGCAGCTA
GGGCTCGCA GCGCGGATCC GAAGACGGT CTCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCGTCAI

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
7121 AGAACAAAGC TCAAACTCAC TCCAACTAGG CCGCTCGCC AGCTCGACTT TCTCGGCTGG TTCACGGCTG GCTACAGCGG
TCTTGTTCG AGTTTGAGTG AGTTTATCG AGGTATCGC CCGGACCGG TCGACCTGAA CAGGCGGACC AAGTGGCGAC CGATGTGCCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
7201 GGGACAGATT TATCAGACGG TCTCTCATG CCGGCCCGC TGGATCTGT TTTGGCTACT CCTGCTTGT CGAGGGGTAG
CCCTCTGTAA ATAGTGTGCG ACAGAGTAGG GCGCGGGGG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCAIC

+2 G I Y L L P N R
7281 GGATCTACCT CCTCCCGAAC CGATGAAGT TGGGTAAAC ACTCCGGCT AAAAAAAA AAAATCTAG AAAGCGCGC
CGTAGATGGA GGAGGGTTG GCTACTTCA ACCCATTTG TGAGGCGGA TTTTTTTT TTTTAGATC TTTCCGGCG

BamHI MluI

7361 CAAGATATCA AGGATCACT ACGGTTAGA GTCGCTGAT CAGCTCGAC TGTGGCTTCT AGTGGCAGC CATCTGTTGT
GTTCTATAGT TCCTAGTGA TGGCAATCT CGAGCGACTA GTCCGAGTG ACACGGAAGA TCAACGGTCG GTAGACAACA

7441 TTGCCCTCG CCGTGGCTT CCTTGACCT GGAAGTGGC ACTCCCACTG TCCTTTCTTA ATAAATGAG GAAATTGAT
AACGGGAGG GGGCAGGAA GGAACCTGGA CCTTCCAGG TGAGGGTAG AGGAAAGAT TATTTACTC CTTTACGTA

pCMV-delINS35

| | |
|------|---|
| 7521 | CGCATTGTCT GAGTAGGTGT CATTCTATTTC TGGGGGGTGG CAGACCAAGG GGGAGGATTG GGAAGACAAAT GCGTAACAGA CTCATCCAGA GTAAGATAAG ACCCCCCACC CCACCCGGTC CTGTGGTTCC CCCTCTTAAC CCTTCTGTAA |
| 7601 | AGCAGGGATG CTGGGGAGCT CTTCGCTTC CTGCTCACT GACTGCTGC GTCGTGTGT TCGGTGGGG CGAGGGGTAT TCGTCCGTAC GACCCCTGCA GAAGGGGAAG GAGCGATGA CTGAGCGAG CGAGCCAGCA AGCCGACGCC GCTCGCCATA |
| 7681 | CAGCTCACTC AAAGGGGTA ATAGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAGA ACATGTGAGC AAAAGGCCAG GTGCGATGAG TTTCGCCCAT TATGCCAATA GGTCTCTTAG TCCCTATTG CGTCTTTCT TGTACTACG TTTTCGGGTC |
| 7761 | CAAAAAGGCCA GGAACCGTAA AAAGGCCGGG TTGCTGGCT TTTTCCATAG GCTCGGCCCC CTGACGAGC ATCACAATAA GTTTTCGGGT CTTTGGGATT TTTCCGGGCG AACGACCGCA AAAAGGTATC CGAGCGGGGG GGAACCTTGG AGGGAGCAGC |
| 7841 | TCGACGGCTCA AGTCAGAGGT GGCGAACCC GACAGACTA TAAAGTACC AGCGTTTCC CCCTGGAAGC TCCTCGTTCG AGCTCGAGT TCACTCTCCA CCGCTTTGGG CTGTCTGTAT ATTCTATGG TCCGCAAGG GGAACCTTGG AGGGAGCAGC |
| 7921 | GCTCTCTGT TCCGACCCCTG CCGCTTACCG GATACCTGTC CCGCTTTCTC CTTTCGGGAA GCGTGGCGT TTCTCAATGC CGAGAGGACA AGGCTGGGAC GCGCAATGGC CTATGGACAG GGGGAAGAG GGAAGCCCTT CGCACCCGGA AAGATTAGC |
| 8001 | TCACGGCTGA GGTATCTCAG TTCCGTTAG GTCTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCGCG TTGAGCCGGA ACTGCGACAT CCATAGAGTC AAGCCATC CAGCAAGTC GGTTCGACC GTCACAGTG CTTGGGGGGC AAGTCGGGCT |
| 8081 | CGGCTGGGC TTATCCGTA ACTATCTGT TGACTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GGCAGCGGG AATAGGCCAT TCATAGCAGA ACTCAGGTTG GGGCATTTCTG TGCTGAATAG CCGTGACCGT CGTGCGTGAC |
| 8161 | GTAACAGGT TACGAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGTGGC CTAACCTAGG CTACACTAGA CACTTCCCTA ATGCTCTCGC TCCATATGC GGGCAGCATG TCTCAAGAAG TTCACCAAG CATTGATGCC GATGTCTATC |

FIG. 5-Page 14

pCMV-deINS35

8241 AGCAGATAT TTGCTATCTG CGCTCTGCTG AAGCCAGTTA CTTTCGGAAA AAGAGTTGGT AGCTCTTGAT CGGCAACA
TCCTGTCTA ACCATAGAC GCGACACAC TTCGGTCAAT GGAAGCTTT TTCTCACCA TCGAGAATA GCGGTTGT

8321 AACCCCGCT GTAGCGGTG GTTTTTTGT TTGCAAGCAG CAGATTAGCG CGAGAAAAA AGGATCTCAA GAAATCCTT
TTGGTGGGA CCAATGGCAC CAAAAAACA AAGTTGCTC GTCTAATGCG CGTCTTTTT TCTTAGAGTT CTCTAGGAA

8401 TGATCTTTTC TAGCGGGTCT GACGCTCACT GGAAGAAAA CTCAGTTAA GGGATTTGG TCAATGAGTT ATCAAAAAGG
ACTAGAAAAG ATGCCCCAGA CTCGAGTCA CTTGCTTTT GAGTGCAATT CCTAAAAACC AGTACTCTAA TAGTTTTTC

8481 ATCTTCACT AGATCCTTTT AAATTAAAAA TGAATTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTGTGACAG
TAGAGTGA TCTAGAAAAA TTAAATTTTT ACTTCAAAAT TTAGTTAGT TTCAATATA CTCATTTGAA CCAGCTGTC

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCACCGATC TGTCTATTTC GTTATCCAT AGTTGCCCTGA CTCGCCCTGC
AATGGTTAGC AATTAGTAC TCGGTGGATA GAGTCGCTAG ACAGATAAG CAAGTAGGTA TCAACGGACT GAGGGGAGC

8641 TGTAGATAAC TAGGATAGCG GAGGGCTTAC CATCTGGCC CAGTGTGCA ATGATACCG GAGACCCAGC CTCACCGCT
ACATCTATTG ATGCTATGCC CTCGCCAATG GTAGACGGG GTACAGAGCT TACTATGCG CTCGTGGTGC GAGTGGCCGA

8721 CCAGATTAT CAGCAATAA CCAGCCAGCC GGAAGGCGG AGCGAGAAAG TGGTCTGCA ACTTTATCCG COTCATCCA
GCTCTAATA GTGTTATT GTCTGCTGG CCTTCCGCGC TCGCTCTTC ACCAGGAGCT TGAATAGGC GGAGGTAGT

8801 GTCTATTAT TGTGCCCGG AAGCTAGAT AAGTAGTTG COAGTTAATA GTTTGCCCAA CTTTGTGTC ATTGCTACAG
CAGATAATA ACAACGGCC TTGATCTCA TTCAATCAAG GCTCAATT CAAAGCGCTT GCAACAACGG TAAGGTATG

8881 GCATGCTGTG GTACGCTCG TCGTTGGTA TGGTTCAAT CAGCTCGGT TCCCAACGAT CAAGGCGAGT TACATATCC
CGTAGACCA CAGTGGGAG ACCTAACCAT ACCGAAGTAA GTGAGGCCA AGGGTCTCTA GTTCCGCTCA ATGTACTAG

FIG. 5-Page 15

pCMV-deINS35

8961 - CCCATGTTGT GCAAAAAGC GGTATGCTC TTCGGTCTC GCATCGTTGT CAGAAGTAAG TTGCGCCGAG TGTATCACT
GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTATCAACA GTCTTCATT AACCGCGTC ACAATAGTA

9041 CATGGTTATG CGAGCACTGC ATAATTCTCT TACTGTGATG CCATCCGTAA GATGCTTTTC TCTGACTGCT GAGTACTCAA
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCAIT CTACGNAAG ACACGTACCA CTCATGAGTT

9121 CCAAGTCATT CTCGCAATAG TGTATGCGC GACCGAGTTG CTCCTTGCCG GGTCAATAC GGGATAATAC CGCGCCACAT
GGTTCAGTAA GACTCTTAIC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTATG CCTATTATG GCGCGGTGA

9201 AGCAGAACTT TAAAGTGTCT CATCATTTGA AAAGCTTCTT CGGGCGGAAA ACTCTCAAGG ATCTTACCGC TGTTCAGATC
TCGCTCTGAA ATTTTCACGA CTAGTAACCT TTTCGAAGAA GCGCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTGATG TAACCCACTC GTGACCCCAA CTGATCTTCA GCACTCTTTA CTTTACCAG GCTTTCTGGG TGACCAAAA
GTCAAGTAC ATTGGGTGAG CACGTGGGTT GACTAGAGT CCGTAAAAAT GAAAGTGCTC CCAAGACCC ACTCGTTTTT

9361 CAGGAAGGCA AATCCCGCA AAAAAAGGAA TAAGGGGAC AGGAAATGT TGAATACTCA TACTCTTCTT TTTTCAATAT
GTCTTTCCGT TTACGGGTG TTTTCCGCTT ATTCCGGTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTTATA

9441 TATTGAAGA TTTATCAGG TTATTGCTC ATGACCGGAT ACATATTGA ATGATTTAG AAAATAAAC AAATAGGGT
ATAACTTGT AAATAGTCCC AATAACAG TACTCGCTA TGTATAACT TACATAAATC TTTTATTG TTTATCCCCA

9521 TCCGCCACA TTTCCCGCA AAGTGGACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGG
AGGGGGGTG AAGGGGCTT TTACCGGTG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601 GTATCAGGAG GCGCTTCTG C
CATAGTCTC CGGGAAGCA G

FIG. 5-Page 16

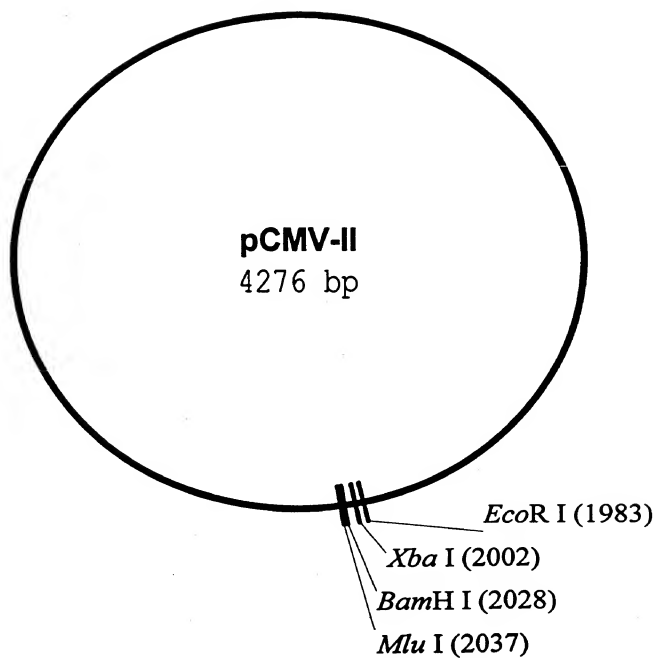


FIG. 6

pcMV-II

| | |
|-----|--|
| 1 | TCGCGGGTTT CGGTGATGAC GGTCAAAACC TCTGACACAT GCAGTCCCG GAGCGGTCA CAGCTTGTCT GTAAGCGAT AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGA CGTCAGGCG CTTGCCAGT GTCGAACAGA CATTGGCGTA |
| 81 | GC CGGAGCA GACAAGCCG TCAGGCGGGT TCAGCGGGT TTGGCGGGT TCGGGGCTG CTTAACTATG CGGCATCAGA CGGCGCTCGT CTGTTGGCG AGTCCGGCG AGTCGCCAC AACGCGCCAC AGCGCGGACC GAATTGATAC GCGTAGTCT |
| 161 | GCAGATTGTA CTGAGATGC ACCATAGAA GCTTTTGG AAGGCTAG CCTCGAAAA AGCCTCCTCA CTACTTCTGG GGTCTAACAT GACTCTCAG TGGTATAC TTGAAAAAGT TTTGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC |
| 241 | AATAGCTCAG AGGCGGAGG GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCA TGGGGCGGAG AATGGCGCGA TTATCGAGTC TCGGGCTCG CGGAGCGCG AGACGTATT ATTTTTTTA ATCAGTCGGT ACCCGGCTC TTACCGGCT |
| 321 | ACTGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAAT ACCTTCTATC TATATCATA TATGTACATT TATATTGGCT TGACCGGCCC CTCCTTAAT AACCGATAAC CGGTAACGTA TGAACATAG ATATAGTATT ATACATGTAA ATATACCGA |
| 401 | CATGTCCAAT ATGACGCCCA TGTGACATT GATTATTGAC TAGTATTAA TAGTAATCAA TTACGGGCTC ATTAGTTCAAT GTACAGGTTA TACTGGGGT ACACGTGTA CTAATAACTG ATCAATATT ATCATTAGTT AATGCCCCAG TAATCAAGTA |
| 481 | AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCGCGC TGGCTGACCG CCCAACGACC CCGGCCCAAT TCGGGTATAT ACCTCAAGC GCAATGATT GAATGCCATT TACCGGGCGG ACGGACTGCG GGGTTGCTGG GGGCGGGTAA |
| 561 | GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTAGCTCA ATGGGTGGAG TATTATCGGT CTGCAAGTTAT TACTGCATAC AAGGTATCA TTGGGTAT CCCTGAAAGG TAACTGCAAT TACCCACCTC ATAAATGCCA |
| 641 | AAACTGCCA CTTCGCAGTA CATCAAGTGT ATCATATGCC AAGTCGCCC CCTATTGACG TCAATGACGG TAAATGGGCC TTTGACGGGT GAACGGTCAAT GTAGTTTACA TAGTATACGG TTCAGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG |

FIG. 7-Page 1

pCMV-II

| | |
|------|--|
| 721 | CGCTGGGATT ATGCCAGTA CATGACCTTA CGGACCTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCCTATTAC CGGACGGTAA TACGGGTCACT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATATCACT AGGATAATG |
| 801 | CATGTGTGAT CGGTTTTGGC AGTACACCAA TGGCGCTGGA TAGCGTTTG ACTCACGGGG ATTTCGAAGT CTCACCCCA GTACACTTAC GCCAAAACCG TCATGTGTT ACCCGCACT ATCCCAAC TGAGTCCOC TAAAGTTCA GAGTGGGCT |
| 881 | TTGACCTCAA TGGGAGTTTG TTTTGGCAGC AAAATCAAG GGACTTTTCA AAATGTCGA ATAAACCCCG CCGCTTGAG AACTCGAGTT ACCCTCAAC AAAACCTGG TTTAGTTGC CCTGAAAGT TTTACAGCAT TATTGGGGG GGGCAACTGC |
| 961 | CAAAATGGCG GTAGCGTGT AGGCTGGGAG GTCTATATAA CAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGGAGAG CTTTACCCGC CATCGGACA TGGCACCCCTC CAGATATATT CGTCTCGAGC AAATCACTTG CCACTCTAGC GGACCTCTGC |
| 1041 | CGATCCAGCG TCTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGGGCGG GGAACGCTGC ATTGGAACCG GGTAGTGGC ACAAAACTGG AGGTAICTTC TGTGGCCCTG CGTAGTCCGG AGGGCGGGC CCTTGGCAGC TAACTTGGC |
| 1121 | GGATTCCCGC TCCCAAGAGT GACGTAAGTA CCGGCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAT GCATGCTATA CCTAAGGGGC AGGTTCTCA CTGCATTCAI GCGGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT |
| 1201 | CTGTTTTTGG CTGGGGGCT ATACACCCCG GCTCCTTATG CTATAGTGA TGCTATAGCT TAGCTATAG CTGTGGGTTA GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT |
| 1281 | TTGACCAITTA TTGACATC CCCTATTGTT GACGATACCT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT AACTGTAAT AACTGTCAG GGGATACCA CTGCTATCAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTCTTGATA |
| 1361 | CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGATGCG GTCCATTAT GAGATAACCG ATATACGTT ATGAGACAGG AAGTCTCTGA CTGTGGCTGA GACATAAATA TGTCTACCC CAGGTAAATA |

FIG. 7 - page 2

pCMV-II

1461 TATTTACAAA TTCAATATA CAACAAGCC GTCCCCCGTC CCCCACTTT TTATTAAACA TAGCTGGGA TCTCCGACAT
ATAAATGTTT AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGGTGAAA AATAATTTGT ATCCACCCCT AGAGGTGTGA

1521 CTCGGGTACG TGTTCGGGAC ATGGGCTCTT CTCGGGTAGC GCGGAGCTT CCACATCGGA GCCGTGTGC CATCGGTCCA
GAGCCCATGC ACAAGGCTG TACCGGAGAA GAGGCCATCG CCGGCTCGAA GGTGTAGGCT CGGACACAGG GTAGGAGGT

1601 GCGGCTCATG GTCGTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCAC
CGCCGAGTAC CAGCGAGCGG TCGAGGAACG AGGATTCTCA CTTCCGCTCT CAATCCGTGT GGTGTTACGG GTGGTGGTGG

1681 AGTGTGCGGC ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAAATGAGT CGGAGATTGG CTTGGCACCT GGACGCGAT
TCACACGGCG TTTTCGGGCA CGGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCCTGGA CTTCCGCTCA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGCGAGCT CAGTTCTTGT ATTCTGATAA GAGTCAAGGG TAACTCCGCT
CCTTCTGAAT TCCGTGCGCG TCTTCTTCTA CGTCCGTGGA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA CTACTGTTTC CTGCGCGCGC CGCCACCAGA CATAATAGCT
AGCCACAGAC AATTGCCACC TCCGCTCACA TCAGACTCCT CATGAGCAAC GAGCGCGCGC CCGGTGTGT GTATTATCGA

1921 GACAGACTAA CAGACTGTC CTTTCCATGG GTCTTTTCTG CAGTCAACCT CCGCGACCTA AGAATTGAGA CTGAGCAGAG
GTGCTGATTT GTCTGACAAG GAAAGGTACC CAGAAAGAC GTCACTGGGA CGAGCTGGAT TCTTAAAGTCT GAGTCTGTTT

XbaI BamHI MluI

2001 TCTAGAAGG GCGCGCAAGA TATCAGGAT CCATACGGG TTAGAGTCGG CTGATCAGCC TCGACTGTGC CTTCAGTTGG
AGATCTTTCC GCGCGGTTCT ATAGTTCTTA GGTGATGCGC AATCGGCGC GACTAGTGG AGCTGACAGG GAAGATCAAC

FIG. 7-Page 3

CMV-II

2081 CCAGCACAAT GTTGTGTC CCTCCCGCT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCTTT TCCTAATAAA
GCTCGGTAGA CAACAACGG GAGGGGGA CCGAAGGAAC TGGGACCTTC CACGGTGAG GTGACAGGAA AGCAITATTT

2161 ATGAGGAAT TGCATCGCAT TGTCTGAGTA GGTGTCAATC TATCTGGG GGTGGGTGG GGCAGGACAG CAAGGGGGAG
TACTCCITTA AGCTAGGCTA ACAGACTCAT CCACAGTAAG ATAGACCCC CCACCCCAAC CCGTCTCTGC GTTCCCGCTC

2241 GATTGGGAAG ACAATAGCAG GCATGCTGGG GAGCTCTTCC GCTTCCTGCG TCACTGACTC GCTCCGCTCG GTCTTCGCG
CTAACCCCTC TGTATCGTC CGTACGACCC CTCGAGAAGG GAAAGGAGCG AGTGACTGAG CGACGGGAGC CAGCAAGCGG

2321 TCGCGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATAGC GTTATCCACA GATCAGGGG ATAAOCGAGG AAGAACAATG
ACGCCGCTCG CCATAGTCCA GTGAGTTTCC GCCATTATGC CAATAGGTCT CTTAGTCCC TATTGGCTCC TTTCTTGTAC

2401 TCAGCAAAAG GCCAGCAAA GGCAGGAAC CCTAAAAGG CCGCTTGTCT GCGTTTTTC CATAGGCTCC GCCCCCTCGA
ACTCGTTTT CCGTCTGTTT CCGGTCTTTC GCATTTTTCC GGCGCAACGA CCGCAAAAG GTATCCGAGG CCGGGGGACT

2481 CGAGCATCAC AAAATCGAC GTCTCACTCA GAGGTGGCGA AACCGACAG GACTATAAAG ATACGAGGG TTTCCGCGTG
GCTCTAGTG TTTTAGCTG CGAGTTCAGT CTCACCGCT TTGGGCTGTC CTGATATTC TATGCTCCG AAAGGGGAC

2561 GAAGTCCCT GGTGGCTCT CTTGTTCGA CCGTGGCGGT TACGGATAC CTGTGGGCT TTCTCGCTTC GGAAGCGGTG
CTTCAGGGA GCACCGGAGA GAACAAGGCT GGGACGGCG ATGGCTATG GACAGGCGGA AAGAGGGAAG CCGTTCGGAC

2641 GCGCTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTGCG TGTAGTCTT TCGTCCAAAG CTGGGCTGTG TGCACGAAC
CGCGAAAGG TTACGACTGC GACATCCATA GAGTCAAGCC ACATCCAGGA AGCGAGTTTC GACCCGACAC ACGTGTGTTG

2721 CCCCCTTCAG CCGGACCGCT GGGCTTATC CGGTAACATN CGTCTTGACT CCAACCCGCT AAGACAGCAG TTATCCGCAC
GGGGCAAGTC GGGGTGGCGA CCGGGAATAG GCCATTGATA GCAGAACTCA GTTTGGGCGA TTCTGTCTG AATAGGGGTG

FIG. 7-Page 4

pCMV-II

2801 TGGCAGGAGC CACTGGTAAAC AGGATTAGCA GAGCGAGGTA TGTAGCGGCT GCTACAGACT TCTTGAAGTG GTGGCCTAAC
ACCGTGTCTG GTGACCAATG TCCTAATCGT CTGCTCCAT ACATCGGCCA CGATGTCTCA AGAACTTCAC CACCGGATG

2881 TACGGCTACA CTAGAAGAC AGTATTGGT ATCTGGGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGCTAGTTC
ATCGCGAIGT GATTTCTCTG TCATAAACCA TAGACGGGAG ACGACTCGG TCAATGGAAG CCTTTTCTC AACATCCAG

2961 TTGATCGGGC AAACAAACCA CCGTGTGTAG GGGTGGTTTT TTTGTTTGCA AGCAGGAGAT TAGCGCGAGA AAAAAGGAT
AACTAGGCGG TTGTTTGGT GCGGACCATC CCGACCAAAA AAACAACT TCCTCTCTCA ATGCGGCTCT TTTTTCCTA

3041 CTCAGAAGA TCGTTTGATC TTTTCTACGG GGTCTGACGC TCACTGGAAC GAAAACTCAC GTTAAGGAT TTTGGTCAATG
GAGTTCTTCT AGGAAACTAG AAAAGATGGC CCAGACTGGG AGTCACCTTG CTTTTGAAGT CAATGCCCTA AAACCAAGTAC

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAATT AAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGATTA
TCTAATAGTT TTCTCTAGAA GTGGATCTAG GAAAAATTAA TTTTACTTC AAAATTIAGT TAGATTTCAT ATATCTCAT

3201 AACTTGTCT GACAGTTACG AATCGTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCTGTCA TCCATAGTTG
TTGAACCGA CTGTCAATGG TTAGCAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

3281 CCTGACTCCC CCGTCTAG ATAACTACGA TACGGAGCG CTTACCATCT GCGCCAGTG CTGCAATGAT ACCCGAGAC
GGACTGAGGG CGAGCACATC TATTGATGCT ATGCCCTCCC GAATGTGAGA CCGGGGTAC GAGGTTACTA TGGCGTCTG

3361 CCAGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACGCG CAGCGGGAAG GCGCGAGCGC AGAAGTGCTC CTGCAACTTT
GGTGGAGTG GCCGAGTCT AAATAGTCTG TATTGTGCG GTGGGCTTC CCGGCTCGG TCTTCACCAG GAGTTGAAA

3441 ATCGGCGCTC ATCCAGTCTA TTAATTGTT GCGGGAAGCT AGAGTAAGTA GTTCGGCACT TAATAGTTTG CGCAAGGTTG
TAGCGGAGG TAGGTCAGAT AATTACAAAC GGCGCTTGA TCTGATTTCAT CAACGGGTCA ATTATCAAC GCGTTCGAAC

FIG. 7-Page 5

pCMV-II

3521 TTGGCATTCG TACAGGATC GTGGTGTAC GCTCGTCCT TGGTATGGT TCATTAGCT CCGGTCCCA ACATCAAGG
AAGGTAACG ATGTCCGTAG CACCACAGTG CGACGAGCA ACCATACCGA AGTAAGTGA GGGCAAGGT TGTACTTCC

3601 CCAGTTACAT GATCCCCCAT GTTGTGAAA AAAGCGTTA CCTCTTCGG TCCTCCGATC GTTGTGAGAA GTAAGTTGG
GCTCAATGTA CTAGGGGTA CACACGTTT TTTCCCAAT CGAGGAAGCG AGGAGCTAG CAACGCTT CATTCACCG

3681 CGCAGTGTA TCAGTCATGG TTATGGGAG ACTGCATAA TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA
GGTCACAAT ACTGAGTACC AATACCTCG TCAGGTATTA AGAATATGAC AGTACGGTAG GCATTCTAAG AAAAGACAT

3761 CTGGTGAGTA CTCACCAAG TCATTCTGAG AATAGTGTAT CGCGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT
GACCACTCAT GAGTTGGTTC AGTAAGACT TTATCACATA CGCGCTGGC TCAACGAGAA CGGCGCGCAG TTATGCGCTA

3841 AATACCGCGC CACATAGCAG AACTTTAAA GTGCTCATCA TTGGAAGAAG TTCTTCGGG CGAAACTCT CAAGGATCTT
TTATCGCGCG GTGTATCGTC TTGAAATTTT CACGAGTACT AACCTTTTGC AAGAAGCCCG CTTTTGAGA GTTCTTAGAA

3921 ACCGCTGTTG AGATCCAGTT CGATGTAAC CACTCGTGCA CCGAATGAT CTTCAGCATC TTTTACTTTC ACCAGGTTT
TGCGCACAAC TGTAGGTCAA GCTAGTTG GTGAGCACGT GGGTTGACTA GAAGTGTAG AAAATGAAAG TGGTCGAAA

4001 CTGGGTGAGC AAAACACGGA AGGCAAAATG CCGCAAAAAA GCGAATAAG CCGACACGGA AATCTTCAAT ACTCATACTC
GACCCACTCG TTTTGTGCTT TCGGTTTAC GCGGTTTTT CCGTTATCG CCGTGTCCCT TTACAACCTA TGAGTATAG

4081 TTCTTTTTTC AATATTATG AAGATTAT CAGGTTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA
AAGGAAAAAG TTATAATAC TTGTAATA GTCCCAATA GTCCGACTCT CCGTATCTAT AAATTACAT AAATCTTTT

4161 TAAACAATA GGGGTTCGCG GCACATTTC CCGAAAGTG CCACGTGAGC TCTAAGAAAC CATTTATTAT ATGACATTAA
ATTTCGTTAT CCGCAAGGCG CGCTGAAAGG GCGTTTTTAC GGTGGACTCG AGATTCTTTG GTAATAATAG TACTGTAAT

4241 CCTATAAAA TAGCGGTATC ACGAGGCCCT TTCGTC
GGATATTTT ATCGGCATAG TGCTCCGGGA AAGCAG

FIG. 7 - Page 6

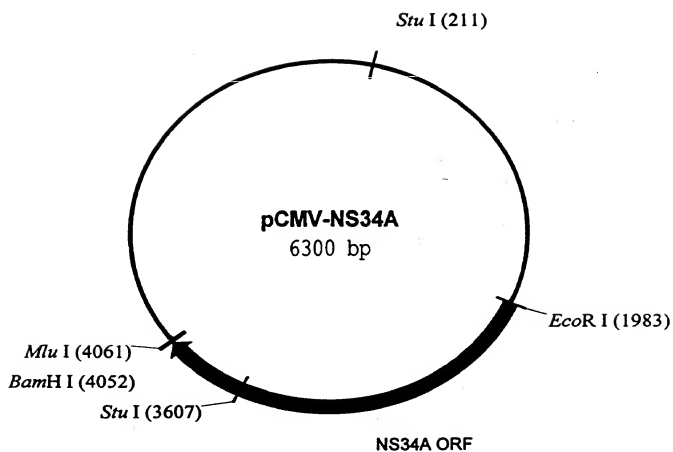


FIG. 8

pCMV-NS34A

| | | | | | |
|-------|------------|-------------|------------|-------------|-------------|
| 1 | TCGCGCGTTT | CGGTGATGAC | GGTGAAAACC | TCTGACACAT | GCAGCTCCCG |
| | AGCGCGCAAA | GCCACTACTG | CCACTTTTGG | AGACTGTGTA | CGTCGAGGGC |
| 51 | GAGACGGTCA | CAGCTTGCT | GTAAGCGGAT | GCCGGGAGCA | GACAAGCCCG |
| | CTCTGCCAGT | GTCGAACAGA | CATTGCGCTA | CGGCCCTCGT | CTGTTGGGGC |
| 101 | TCAGGGCGCG | TCAGCGGGTG | TTGGCGGGTG | TCGGGGCTGG | CTTAACATATG |
| | AGTCCCCGCG | AGTCGCCAC | AACCGCCAC | AGCCCCGACC | GAATTGATAC |
| 151 | CGGCATCAGA | GCAGATTGTA | CTGAGAGTGC | ACCATATGAA | GCTTTTTCGA |
| | GCCGTAGTGT | CGTCTAACAT | GACTCTCAGC | TGGTATACTT | CGAAAAACGT |
| StuI | | | | | |
| ~~~~~ | | | | | |
| 201 | AAAGCCTAGG | CCTCCTCAAAA | AGCCTCCTCA | CTACTTCTGG | AATAGCTCAG |
| | TTTCGGATCC | GGAGGTTTTT | TCGGAGGAGT | GATGAAGACC | TTATCGAGTC |
| 251 | AGGCCGAGGC | GGCCTCGGCC | TCTGCATAAA | TAAAAAAT | TAGTCAGCCA |
| | TCCGGCTCCG | CCGGAGCCGG | AGACGTATTT | ATTTTTTTTA | ATCAGTCGGT |
| 301 | TGGGGCGGAG | AATGGGCGGA | ACTGGGCGGG | GAGGGAATTA | TTGGCTATTG |
| | ACCCCGCCTC | TTACCCGCCCT | TGACCCGCCC | CTCCCTTAAT | AACCGATAAC |
| 351 | GCCATTGCAT | ACGTTGTATC | TATATCATAA | TATGTACATT | TATATTGGCT |
| | CGGTAACGTA | TGCAACATAG | ATATAGTATT | ATACATGTAA | ATATAACCGA |
| 401 | CATGTCCAAT | ATGACCGCCA | TGTTGACATT | GATTATTGAC | TAGTATTAA |
| | TACAGGCTTA | TACTGGCGGT | ACAACGTGTA | CTAATAACTG | ATCAATAATT |
| 451 | TAGTAATCAA | TTACGGGGTC | ATTAGTTTAT | AGCCCATATA | TGGAGTTCGG |
| | ATCATTAGTT | AATGCCCGAG | TAATCAAGTA | TCGGGTATAT | ACCTCAAGGC |
| 501 | CGTTACATAA | CTTACGGTAA | ATGGCCCGCC | TGGCTGACCG | CCCAACGACC |
| | GCAATGTATT | GAATGCCATT | TACCGGGCGG | ACCGACTGCG | GGGTTGCTGG |
| 551 | CCCGCCCAT | GACGTCAATA | ATGACGTATG | TTCCCATAGT | AACGCCAATA |
| | GGCGGGGTAA | CTGCAGTTAT | TACTGCATAC | AAGGGTATCA | TTGCGGTTAT |
| 601 | GGGACTTTCC | ATTGACGTCA | ATGGGTGGAG | TATTACGGT | AAACTGCCCA |
| | CCCTGAAAGG | TAACATGCAGT | TACCCACCTC | ATAAATGCCA | TTTGACGGGT |
| 651 | CTTGGCAGTA | CATCAAGTGT | ATCATATGCC | AAGTCCGCCC | CCTATTGACG |
| | GAACCGTCAT | GTAGTTCACA | TAGTATACGG | TTAGGCGGGG | GGATAACTGC |
| 701 | TCAATGACGG | TAAATGGCCC | GCCTGGCATT | ATGCCCAGTA | CATGACCTTA |
| | AGTTACTGCC | ATTTACCGGG | CGGACCGTAA | TACGGGTCTAT | GTACTGGAAT |
| 751 | CGGACTTTTC | CTACTTGGCA | GTACATCTAC | GTATTAGTCA | TCGTATTAC |
| | GCCCTGAAAG | GATGAACCGT | CATGTAGATG | CATAATCAGT | AGCGATAATG |
| 801 | CATGGTGATG | CGGTTTTTGGC | AGTACACCAA | TGGGCGTGGA | TAGCGGTTTG |
| | GTACCACTAC | GCCAAAACCG | TCATGTGGTT | ACCCGCACCT | ATCGCCAAAC |
| 851 | ACTCAGGGG | ATTTCCAAGT | CTCCACCCCA | TTGAAGTCAA | TGGGAGTTTG |
| | TGAGTGCCCC | TAAAGTTTCA | GAGGTGGGGT | AAGTGCAGTT | ACCTTCAAAC |

pCMV-NS34A

| | |
|------|---|
| 901 | TTTGGCCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG |
| 951 | CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GGCAACTGCG GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT |
| 1001 | GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG |
| 1051 | TGTTTGGACC TCCATAGAAG ACACCCGGAC CGATCCAGCC TCCGCGGGCG ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCCCGGGC |
| 1101 | GGAACGGTGC ATTGGAACGC GGATTCCCGC TGCCAAGAGT GACGTAAGTA CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTTCAT |
| 1151 | CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT |
| 1201 | CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCAAT |
| 1251 | TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC ACCATATCGA ATCGGATATC CACACCAAT AACTGGTAAT AACTGGTGAG |
| 1301 | CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC |
| 1351 | CCACAACAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GGTGTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA |
| 1401 | GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT TATTACAAA CTGTGCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA ATAAATGTTT |
| 1451 | TTACATATA CAACAACGCC GTCCCGGTG CCCGCAGTTT TTATTAAACA AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA ATAAATTTGT |
| 1501 | TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCGGGAC ATGGGCTCTT ATCGCACCTT AGAGGCTGTA GAGCCCATGC ACAAGGCCGT TACCCGAGAA |
| 1551 | CTCCGGTAGC GGCAGGAGTT CCACATCCGA GCCCTGGTCC CATCCGTCCA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT |
| 1601 | GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT |
| 1651 | CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG TGTTCGGCA |
| 1701 | GGCGGTAGGG TATGTGCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA |
| 1751 | GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGT CGAGGCAGCT CCTGGCTCTA CCTTCTGAAT TCCGTGCGCG TCTTCTCTA TCGCCGTGCA |
| 1801 | GAGTGTGTG ATTCTGATAA GAGTCAGAGG TAACTCCCGT GCGGGTGTG CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC |

FIG. 9-Page 2

pCMV-NS34A

| | | | | | |
|------|-------------|------------|------------|-------------|-------------|
| 1851 | TTAACGGTGG | AGGGCAGTGT | AGTCTGAGCA | GTACTCGTTG | CTGCCGCGCG |
| | AATTGCCACC | TCCCGTCACA | TCAGACTCGT | CATGAGCAAC | GACGGCGCGC |
| 1901 | CGCCACCAGA | CATAATAGCT | GACAGACTAA | CAGACTGTTC | CTTTCCATGG |
| | GCGGTGGTCT | GTATTATCGA | CTGTCTGATT | GTCTGACAAG | GAAAGGTACC |
| +2 | | | | EcoRI | M A P |
| | | | | ----- | |
| 1951 | GTCITTTTCTG | CAGTCACCGT | CGTCGACCTA | AGAATTCACC | ATGGCGCCCCA |
| | CAGAAAAGAC | GTCAGTGGCA | GCAGCTGGAT | TCITTAAGTGG | TACCGCGGGT |
| +2 | I T A Y | A Q Q | T R G L | L G C | I I T |
| 2001 | TCACGGCGTA | CGCCAGCAG | ACAAGGGGCC | TCCTAGGGTG | CATAATCACC |
| | AGTGCCGCAT | GCGGGTCGTC | TGTTCCCCGG | AGGATCCAC | GTATTATGG |
| +2 | S L T G | R D K | N Q V | E G E V | Q I V |
| 2051 | AGCCTAACTG | GCCGGGACAA | AAACCAAGTG | GAGGGTGAGG | TCCAGATTGT |
| | TCGGATTGAC | CGGCCCTGTT | TTTGGTTTAC | CTCCCACTCC | AGGTCTAACCA |
| +2 | S T A | A Q T F | L A T | C I N | G V C |
| 2101 | GTCAACTGCT | GCCCAAACCT | TCCTGGCAAC | GTGCATCAAT | GGGGTGTGCT |
| | CAGTTGAGGA | GCGGTTTGGA | AGGACCGTTG | CACGTAGTTA | CCCCACACGA |
| +2 | W T V Y | H G A | G T R T | I A S | P K G |
| 2151 | GGAAGTGTCTA | CCACGGGGCC | GGAACGAGGA | CCATCGCGTC | ACCCAAGGGT |
| | CTGACAGAT | GGTGCCCCCG | CCTTGCTCT | GGTAGCGCAG | TGGGTTCCCA |
| -2 | P V I Q | M Y T | N V D | Q D L V | G W P |
| 2201 | CCTGTCACTCC | AGATGTATAC | CAATGTAGAC | CAAGACCTTG | TGGGCTGGCC |
| | GGACAGTAGG | TCTACATATG | GTTACATCTG | GTTCTGGAAC | ACCCGACCGG |
| +2 | A S Q | G T R S | L T P | C T C | G S S |
| 2251 | CGCTTCGCAA | GGTACCGCT | CATTGACACC | CTGCACTTGC | GGCTCTCTCG |
| | GCGAAGCGTT | CCATGGGCGA | GTAAGTGTGG | GACGTGAACG | CCGAGGAGCC |
| +2 | D L Y L | V T R | H A D V | I P V | R R R |
| 2301 | ACCTTTACCT | GGTCACGAGG | CACGCCGATG | TCATTCCCGT | GCGCGGGCGG |
| | TGGAAATGGA | CCAGTGCTCC | GTGCGGCTAC | AGTAAGGCGA | CGGCGGGCGC |
| +2 | G D S R | G S L | L S P | R P I S | Y L K |
| 2351 | GGTGATAGCA | GGGGCAGCCT | GCTGTGCCCC | CGGCCCAATT | CCCTAGTGAA |
| | CCACTATCGT | CCCCGTCGGA | CGACAGCGGG | GCCGGGTAAA | GGATGAACCT |
| +2 | G S S | G G P L | L C P | A G H | A V G |
| 2401 | AGGCTCTCTG | GGGGGTCCCG | TGTTGTGCCC | CGCGGGGCAC | GCGGTGGGCA |
| | TCCGAGGAGC | CCCCCAGGCG | ACAACACGGG | GCGCCCCGTG | CGGCACCCGT |
| +2 | I F R A | A V C | T R G V | A K A | V D F |
| 2451 | TATTTAGGGC | CGCGGTGTGC | ACCGGTGGAG | TGGCTAAGGC | GGTGGACTTT |
| | ATAAATCCCG | GCGCCACACG | TGGGCACCTC | ACCGATTCCG | CCACCTGAAG |
| +2 | I P V E | N L E | T T M | R S P V | F T D |
| 2501 | ATCCCTGTGG | AGAACCTAGA | GACAACCATG | AGGTCCCCGG | TGTTACACGA |
| | TAGGGACACC | TCTTGGATCT | CTGTTGGTAC | TCCAGGGGCC | ACAAGTGCCT |

FIG. 9-Page 3

pCMV-NS34A

| | | |
|------|--|--|
| +2 | N S S P P V V P Q S F Q V A H L | |
| 2551 | TAACTCCTCT CCACCACTAG TGCCCCAGAG CTTCACGGTG GCTCACCTCC | |
| | ATTGAGGAGA GGTGGTCTAT ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG | |
| +2 | H A P T G S G K S T K V P A A Y A | |
| 2601 | ATGCTCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA | |
| | TACGAGGGTG TCCGTCGCCG TTTTCGTGTT TCCAGGGCCG ACGTATACGT | |
| +2 | A Q G Y K V L V L N P S V A A T L | |
| 2651 | GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT | |
| | CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA | |
| +2 | G F G A Y M S K A H G I D P N I | |
| 2701 | GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA | |
| | CCCGAAACCA CGAATGTACA GGTTCCGAGT ACCTAGCTA GGATTGTAGT | |
| +2 | R T G V R T I T T G S P I T Y S T | |
| 2751 | GGACCGGGGT GAGAACAAAT ACCACTGGCA GCCCATCAC GTACTCCACC | |
| | CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG | |
| +2 | Y G K F L A D G G C S G G A Y D I | |
| 2801 | TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGGC CTTATGACAT | |
| | ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCCGC GAATACTGTA | |
| +2 | I I C D E C H S S T D A T S I L G | |
| 2851 | AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA | |
| | TTATTAAACA CTGCTCACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT | |
| +2 | I G T V L D Q A E T A G A R L V V | |
| 2901 | TTGGCACTGT CTTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG | |
| | AACCGTGACA GGAAGTGGTT CGTCTCTGAC GCCCCCGCTC TGACCAACAC | |
| +2 | L A T A T P P G S V T V P H P N I | |
| 2951 | CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT | |
| | GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA | |
| +2 | E E V A L S T T G E I P F Y G K | |
| 3001 | CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG | |
| | CGCTCTCCAA CGAGACAGGT GGTGGCCTCT CTAGGGAAAA ATGCCGTTC | |
| +2 | A I P L E V I K G G R H L I F C H | |
| 3051 | CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTCTGTGCAT | |
| | GATAGGGGGA GCTTCATTAG TTCCCCCCTC CTGTAGAGTA GAAGACAGTA | |
| +2 | S K K K C D E L A A K L V A L G I | |
| 3101 | TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTCG CATTGGGCTG | |
| | AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA | |
| +2 | N A V A Y Y R G L D V S V I P T | |
| 3151 | CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATGCCGACCA | |
| | GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGCTTGGT | |
| +2 | S G D V V V V A T D A L M T G Y T | |
| 3201 | GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC | |
| | CGCCGTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG | |

FIG. 9-Page 4

pCMV-NS34A

+2 G D F D S V I D C N T C V T Q T V
3251 GGCGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT
CCGCTGAAGC TGAGCCACTA TCTGACGTGA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P
3301 CGATTTTCAGC CTTGACCCCTA CCTTCACCAT TGAGACAACT ACGCTCCCCC
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCAGGGGGG

+2 Q D A V S R T Q R R G E R T G R G K
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGGAG
TTCTACGAGA GAGGGCGTGA GTTGACAGCC CGTCTGACC GTCCCGGGTC

+2 P G I Y R F V A P G E R P S G M F
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCT CCGGCATGTT
GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2 D S S V L C E C Y D A G C A W Y
3451 CGACTCGTCC GTCTCTGTG AGTGTATGA CGCAGGCTGT GCTTGGTATG
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGCG TACGAGCGTA CATGAACACC
TCGAGTGGCG GCGGCTCTGA TGTCATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F
3551 CGGGGCGTTC CCGTGTGCCA GGACATCTTT GAATTTTGGG AGGGCGTCTT
GGCCCCGAGG GGCACACGGT CCTGGTAGAA CTTAAAACCC TCCCGCGTAA

+2 T G L T H I D A H F L S Q T K Q
StuI

3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCAG ACAAGCAGA
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTCGTCT

+2 S G E N L P Y L V A Y Q A T V C A
3651 GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT
CACCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACCGCA

+2 R A Q A P P P S W D Q M W K C L I
3701 AGGGCTCAAG CCCCTCCCC ATCGTGGGAC CAGATGTGGA AGTGTTTGAT
TCCCGAGTTC GGGGAGGGGG TAGCACCCCT GTCTACACCT TCACAACTA

+2 R L K P T L H G P T P L L Y R L
3751 TCGCCTCAAG CCCACCTCC ATGGGCCAAC ACCCTGCTA TACAGACTGG
AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGACGAT ATGTCTGACC

+2 G A V Q N E I T L T H P V T K Y I
3801 CGCCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCACTG GTTTTGTAG

+2 M T C M S A D L E V V T S T W V L
3851 ATGACATGCA TGTGCGCCGA CCTGGAGGTC GTACAGGACA CCTGGGTGCT
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

+2 V G G V L A A L A A Y C L S T G
3901 CGTTGGCGCG GTCTCGGCTG CTTTGGCCGC GTATTGCTG TCACAGGCT
GCAACCGCGC CAGGACCGAC GAAACCGCGC CATACCGGAC AGTTGTCCGA

pCMV-NS34A

| | | | | | | | | | | | | | | | | | | |
|------|----|-----------|----------|----------|---------|---------|---------|---------|----------|----------|--------|------|---------|--------|---------|------|---------|-------|
| | +2 | C | V | V | I | V | G | R | V | V | L | S | G | K | P | A | I | I |
| 3951 | | GCGTGGT | CAT | AGTGGGC | CAGG | GTCGTCT | TGT | CCGGAAG | CC | CGGAAGCC | GGCAAT | CATA | CGC | CAAGTA | TCACCCG | TCC | CAGCAGA | CA |
| | | CGC | CAAGTA | TCACCCG | TCC | CAGCAGA | CA | GGCCCTT | CGG | CCGT | TAGTAT | | | | | | | |
| | +2 | P | D | R | E | V | L | Y | R | E | F | D | E | M | E | E | C | |
| 4001 | | CCTGACAGG | GAAGTCCT | CTA | CCGAGAG | TTC | CGAGAGT | TGG | GATGAGAT | TGG | AAGAGT | GCTA | GGACTGT | CCCC | TTCAGG | AGAT | TTCTC | ACGAT |
| | | GGACTGT | CCCC | TTCAGG | AGAT | GGCTCT | CAAG | CTACTCT | AACC | TTCTC | ACGAT | | | | | | | |
| | | BamHI | | MluI | | | | | | | | | | | | | | |
| | | GGATCC | ACTA | CGCGT | TAGAG | CTCGCT | GATC | AGCCT | CGACT | GTGCCT | TCTA | | | | | | | |
| 4051 | | GGATCC | ACTA | CGCGT | TAGAG | CTCGCT | GATC | AGCCT | CGACT | GTGCCT | TCTA | | | | | | | |
| | | CCTAGT | GTAT | GCGCA | ATCTC | GAGCG | ACTAG | TCGGAG | CTGA | CACGGA | AAGAT | | | | | | | |
| 4101 | | GTGCCAG | CC | ATCTGTT | GT | TGCCCC | TCCC | CCGTGC | CTTC | CTTGACC | CTG | | | | | | | |
| | | CAACGG | TCGG | TAGACA | AACAA | ACGGGG | GAGGG | GGCACG | GGAAG | GAAC | TGGGAC | | | | | | | |
| 4151 | | GAAGGTG | CCA | CTCCCA | CTGT | CCTTTC | TCTAA | TAAATG | AGG | AAATG | CATC | | | | | | | |
| | | CTTCCAC | GGT | GAGGGT | GACA | GGAAAG | GATT | ATTTACT | C | TTTAAC | G | | | | | | | |
| 4201 | | GCATTGT | CTG | AGTAGT | GTG | ATTCTAT | TCT | GGGGGT | GGG | GTGGGG | CAGG | | | | | | | |
| | | CGTAAC | AGAC | TCATCC | ACAG | TAAGATA | AAGA | CCCCCA | CACC | CCCCG | TCC | | | | | | | |
| 4251 | | ACAGCA | AGGG | GGAGG | ATTGG | GAAGACA | ATA | GCAGGC | ATGC | TGGGG | AGCTC | | | | | | | |
| | | TGTCTGT | CCCC | CCTCCT | AACC | CTTCTGT | TAT | CGTCCG | TACG | ACCCCT | CGAG | | | | | | | |
| 4301 | | TTCCGCT | TCC | TCGCTCA | CTG | ACTCGCT | GCG | CTCGGT | CGTT | CGGCTG | CGGC | | | | | | | |
| | | AAGCGA | AAGG | AGCGAG | TGAC | TGAGCG | ACGC | GAGCC | AGCAA | GCCGAC | CGC | | | | | | | |
| 4351 | | GAGCGGT | ATC | AGCTCA | CTCA | AAGCGGT | ATA | TACGGT | TATC | CACAGA | ATCA | | | | | | | |
| | | CTGCCA | CATAG | TCGAGT | GAGT | TTCCGCC | CATT | ATGCCA | CATAG | GTGCTT | AGT | | | | | | | |
| 4401 | | GGGGATA | ACG | CAGGAA | AGAA | CATGTAG | CA | AAAGGCC | AGC | AAAAGG | CCAG | | | | | | | |
| | | CCCTATT | GC | GTCCTT | TCT | GTACACT | CGT | TTTCCG | GTCG | TTTTCC | GGTC | | | | | | | |
| 4451 | | GAACCGT | AAA | AAGGCCG | CGT | TGCTGGC | CGT | TTTCCAT | AGG | CTCCG | CCCCC | | | | | | | |
| | | CTTGGC | CATT | TTCCGG | CGCA | ACGAC | CGCA | AAAGGT | ATCC | GAGGC | GGGG | | | | | | | |
| 4501 | | CTGACG | AGCA | TCACAAAA | AT | CGACGCT | CAA | GTCAGAG | GTG | GCGAAA | CCCG | | | | | | | |
| | | GACTGCT | CGT | AGTGTT | TTTTA | GCTGCG | AGTT | CAGTCT | CCAC | CGCTTT | TGGGC | | | | | | | |
| 4551 | | ACAGGAC | TAT | AAAGATA | ACCA | GGCGTTT | CCC | CCTGGA | AGCT | CCCTCG | TGCG | | | | | | | |
| | | TGTCTG | TATA | TTTCTAT | TGTT | CCGCAAA | AGGG | GGACCT | TCGA | GGGAGC | ACGC | | | | | | | |
| 4601 | | CTCTCT | TGTT | CCGACC | TGC | CGCTTAC | CGG | ATACCT | GTCC | GCCTTT | CTCC | | | | | | | |
| | | GAGAGG | ACAA | GGCTGGG | ACG | GCGAAT | TGGCC | TATGG | ACAGG | CGGAA | AAGAGG | | | | | | | |
| 4651 | | CTTCGGG | AAG | CGTGGC | GCTT | TCTCAAT | GCT | CACGCT | GAG | GTATCT | CAGT | | | | | | | |
| | | GAAGCCC | TTC | GCACCG | CGAA | AGAGTT | ACGA | GTGCG | ACATC | CATAG | AGTCA | | | | | | | |
| 4701 | | TCGGTGT | AGG | TCGTT | CGCTC | CAAGCTG | GGGC | TGTGTG | CACG | AACCCCC | CGT | | | | | | | |
| | | AGCCAC | ATCC | AGCAAG | CGGAG | GTTGAC | CCCG | ACACAC | GTGC | TTGGGGG | GCA | | | | | | | |
| 4751 | | TCAGCCC | AGAC | CGCTGC | GCGCT | TATCCG | GATA | CTATCG | TCTT | GATGCT | CAACC | | | | | | | |
| | | AGTCGGG | CTG | GCGACG | CGGA | ATAGG | CCATT | GATAG | ACGAA | CGAGG | TGG | | | | | | | |
| 4801 | | CGGTAAG | ACA | CGACTT | ATCG | CCACTG | GCAG | CAGCCAC | TGG | TAAACAG | GATT | | | | | | | |
| | | GCCATT | TCTG | GCTGA | ATAGC | GGTGAC | CGTC | GTCGGT | GACC | ATTGTC | TCTAA | | | | | | | |

FIG. 9-Page 6

pCMV-NS34A

| | | | | | |
|------|--------------------------|---------------------------|---------------------------|--------------------------|--------------------------|
| 4851 | AGCAGAGCGA TCGTCTCGCT | GGTATGTAGG CCATACATCC | CGGTGCTACA GCCACGATGT | GAGTTCCTGA CTCAAGAACT | AGTGGTGGCC TCACCACCGG |
| 4901 | TAACACGGC ATTGATGCCG | TACACTAGAA ATGTGATCTT | GGACAGTATT CCTGTCAATA | TGGTATCTGC ACCATAGACG | GCTCTGCTGA CGAGACGACT |
| 4951 | AGCCAGTTAC TCGGTCAATG | CTTCGGAAAA GAAGCCTTTT | AGAGTTGGTA TCTCAACCAT | GCTCTTGATC CGAGAAGTAC | CGGCAAAACA GCCGTTTGT |
| 5001 | ACCACCGCTG TGGTGGCGAC | GTAGCGGTGG CATCGCCACC | TTTTTTTGT AAAAAAACAA | TGCAAGCAGC ACGTTTCGTC | AGATTACGCG TCTAATGCGC |
| 5051 | CAGAAAAAAA GTCTTTTTTT | GGATCTCAAG CCTAGAGTTC | AAGATCCTTT TTCTAGGAAA | GATCTTTTCT CTAGAAAAGA | ACGGGGTCTG TGCCCCAGAC |
| 5101 | ACGCTCAGTG TGGGAGTCAC | GAACGAAAAC CTTGCTTTTG | TCACGTTAAG AGTGCATATC | GGATTTTGGT CCTAAAACCA | CATGAGATTA GTACTCTAAT |
| 5151 | TCAAAAAGGA AGTTTTTCT | TCTTCACCTA AGAAGTGGAT | GATCCTTTTA CTAGGAAAAAT | AATTAAAAAT TTAATTTTTA | GAAGTTTAA CTTCAAAAT |
| 5201 | ATCAATCTAA TAGTTAGATT | AGTATATATG TCATATATAC | AGTAACTTG TCATTTGAAC | GTCTGACAGT CAGACTGTCA | TACCAATGCT ATGGTTACGA |
| 5251 | TAATCAGTGA ATTAGTCACT | GGCACCTATC CCGTGGATAG | TCAGCGATCT AGTCGCTAGA | GTCTATTTTC CAGATAAAGC | TTATCCATA AAGTAGGTAT |
| 5301 | GTTGCCTGAC CAACGGACTG | TCCCCGTCGT AGGGGACGCA | GTAGATAACT CATCTATTGA | ACGATACGGG TGCTATGCCC | AGGGCTTACC TCCCAGATGG |
| 5351 | ATCTGGCCCC TAGACCGGGG | AGTGCTGCAA TCACGACGTT | TGATACCGCG ACTATGGCGC | AGACCCACGC TCTGGGTGCG | TCACCGGCTC AGTGGCCGAG |
| 5401 | CAGATTATC GTCTAAATAG | AGCAATAAAC TCGTTATTTC | CAGCCAGCCG GTCGGTCGGC | GAAGGGCCGA CTTCCCGGCT | GCGCAGAAGT CGCGCTTCA |
| 5451 | GGTCCGCAA CCAGGACGTT | CTTTATCCGC GAAATAGGGG | CTCCATCCAG GAGGTAGGTC | TCTATTAATT AGATAATTAA | GTGCGCGGGA CAACGGCCCT |
| 5501 | AGCTAGAGTA TCGATCTCAT | AGTAGTTCGC TCATCAAGCG | CAGTTAATAG GTCAATTATC | TTTGCGCAAC AAACGCGTTG | GTGTTGCCA CAACAACGGT |
| 5551 | TTGCTACAGG AACGATGTCC | CATCGTGGTG GTAGCACCCAC | TCACGCTCGT AGTGCAGCA | CGTTTGGTAT GCAAACCAT | GGCTTCATTC CCGAAGTAAG |
| 5601 | AGCTCCGGTT TCGAGGCCAA | CCCAACGATC GGGTTGCTAG | AAGGCGAGTT TTCCGCTCAA | ACATGATCCC TGTAAGTAGG | CCATGTTGTG GGTACAACAC |
| 5651 | CAAAAAGCG GTTTTTTCG | GTTAGTCTCT CAATCGAGGA | TCGGTCTCCT AGCCAGGAGG | GATCGTTGTC CTAGCAACAG | AGAAGTAAGT TCTTCATTCA |
| 5701 | TGGCCGCAGT ACCGGCGTCA | GTTATCACTC CAATAGTGAG | ATGGTTATGG TACCAATACC | CAGCACTGCA GTCTGTACGT | TAATTCCTCT ATTAGAGAA |
| 5751 | ACTGTATGTC TGACAGTACG | CATCCGTAAG GTAGGCATTC | ATGCTTTTCT TACGAAAAGA | GTGACTGGTG CACTGACCAC | AGTACTCAAC TCATGAGTTG |

pCMV-NS34A

| | | | | | |
|------|-------------|------------|-------------|------------|------------|
| 5801 | CAAGTCATTC | TGAGAATAGT | GTATGCGGCG | ACCGAGTTGC | TCTTGCCCGG |
| | GTTTCAGTAAG | ACTCTTATCA | CATACGCCGC | TGGCTCAACG | AGAACGGGCC |
| 5851 | CGTCAATACG | GGATAATACC | GCGCCACATA | GCAGAACTTT | AAAAGTGCTC |
| | GCAGTTATGC | CCTATTATGG | CGCGGTGTAT | CGTCTTGAAA | TTTTCACGAG |
| 5901 | ATCATTGGAA | AACGTTCTTC | GGGGCGAAAA | CTCTCAAGGA | TCTTACCGCT |
| | TAGTAACCTT | TTGCAAGAAG | CCCCGCTTTT | GAGAGTTCTT | AGAATGGCGA |
| 5951 | GTTGAGATCC | AGTTCGATGT | AACCCACTCG | TGCACCCAAC | TGATCTTCAG |
| | CAACTCTAGG | TCAAGCTACA | TTGGGTGAGC | ACGTGGGTTG | ACTAGAAGTC |
| 6001 | CATCTTTTAC | TTTCACCAGC | GTTTCTGGGT | GAGCAAAAAC | AGGAAGGCAA |
| | GTAGAAAATG | AAAGTGGTCG | CAAAGACCCA | CTCGTTTTTG | TCCTTCCGTT |
| 6051 | AATGCCGCAA | AAAAGGGAAT | AAGGGCGACA | CGGAAATGTT | GAATACTCAT |
| | TTACGGCGTT | TTTCCCTTA | TTCCCGCTGT | GCCTTTACAA | CTTATGAGTA |
| 6101 | ACTCTTCCTT | TTTCAATATT | ATTGAAGCAT | TTATCAGGGT | TATTGTCTCA |
| | TGAGAAGGAA | AAAGTTATAA | TAAC TTCGTA | AATAGTCCCA | ATAACAGAGT |
| 6151 | TGAGCGGATA | CATATTTGAA | TGTATTTAGA | AAAATAAACA | AATAGGGGTT |
| | ACTCGCCTAT | GTATAAACTT | ACATAAATCT | TTTTATTGT | TTATCCCAAA |
| 6201 | CCGCGCACAT | TTCCCCGAAA | AGTGCCACCT | GACGTCTAAG | AAACCATTAT |
| | GGCGCGTGTA | AAGGGGCTTT | TCACGGTGGA | CTGCAGATT | TTTGGTAATA |
| 6251 | TATCATGACA | TTAACCTATA | AAAATAGGCG | TATCACGAGG | CCCTTCGCT |
| | ATAGTACTGT | AATTGGATAT | TTTATCCGC | ATAGTGCTCC | GGGAAAGCAG |

FIG. 9-Page 8

Diagram 1

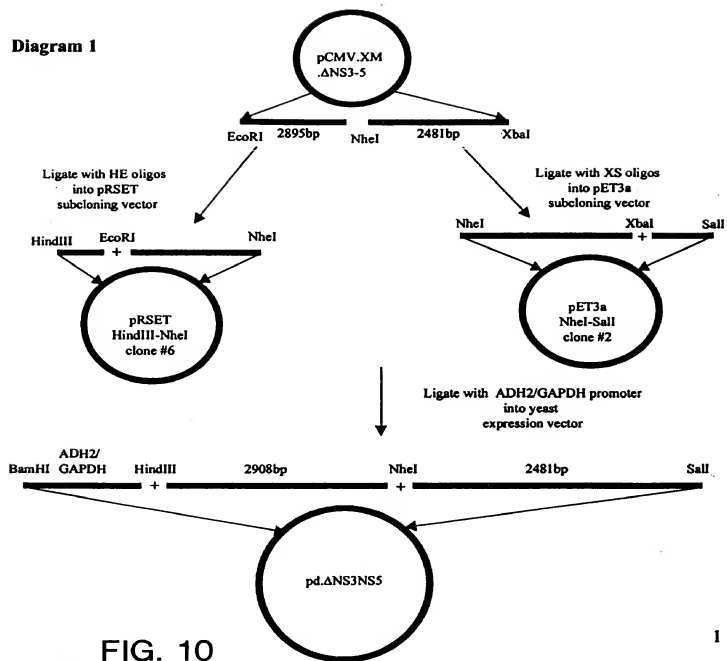


FIG. 10

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal
2 AGCTTACAAAACAAATTCAACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA
TCGAATGTTTTGTTTAAAGTGGTACCGACGTATACGTGAGTCCCGATATCCACGATCAT
^
1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,
^
LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly
62 CTCACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG
GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCAAGATGTACAGGTTCCGAGTAGCC
^
IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr
122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC
TAGCTAGGATTGTAGTCTTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG
^
122 CLAI,
^
SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle
182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATA
AGGTGGATGCGGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGCAATACTGTATTAT
^
IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu
242 ATTTGTGACGAGTGCCTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCTTT
TAACACTGCTCTACGGTGAGGTGCCTACGGTGAGGTAGAACCGTAACCGTGACAGGAA
^
AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly
302 GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC
CTGTTCTGTTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCG
^
309 ALWN1,
^
SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle
362 TCCGTCACTGTGCCCATCCCAACATCGAGGAGTTGCTCTGTCCACCACCGGAGAGATG
AGGCAGTGACACGGGGTAGGGTTGTAGTCTCTCCAACGAGACAGGTGGTGGCCTCTCTAG
^
ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe
422 CCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC
GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG
^
CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn
482 TGCATTCAAAGAAGAAGTGCAGCAACTCGCCGCAAGCTGGTCGATTGGGCATCAAT
ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACACCGGTAACCCGTAGTTA
^
AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal
542 GCCGTGGCCTACTACCGCGTCTTGACGTGTCCTGTCATCCGACCAAGCGCGATGTTGTC
CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGTACAACAG
^
556 SAC2, 566 DRD1,
^
ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp
602 TTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGCGACTTCGACTCGGTGATAGAC
CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTAAGCTGAGCCACTATCTG
^
621 BSPH1,
^
CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

662 TGCAATACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCCTACCTTCACCATTTGAG
ACGTTATGACACAGTGGGCTCTGTACAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACTC

ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg
722 ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG
TGTTAGTGCAGAGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCGCTCCTGACCGTCC

GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp
782 GGGAAAGCCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTTCGAC
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAAGTGC

822 BGLI, 839 DRD1,

SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla
842 TCGTCCGTCTCTGTGAGTGCTATGACGAGGCTGTGCTTGGTATGAGCTCAGCCCGCC
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATCTCGAGTGGGGGCGG

887 SACI,

GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp
902 GAGACTACAGTTAGGCTACGAGCGTACATGAACCCCGGGCTTCCCGTGTGCCAGGAC
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACAGGTCCTG

937 SMAI XMAI,

HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu
962 CATCTTGAATTTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCCACTTTCTA
GTAGAAGTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT

991 STUI,

SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal
1022 TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGTAGCGTACCAAGCCACCGTG
AGGGTCTGTTTCGTCTCACCCCTCTTGGAAAGGAATGSAACCATCGCATGGTTCGGTGGCAC

1075 DRA3,

CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg
1082 TCGCGTAGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCCG
ACCGCATCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAAACTAAGCG

LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn
1142 CTCAGCCACCCCTCCATGGGCCAACACCCCTGCTATACAGATGGGGCGCTGTTTCAAGT
GAGTTCGGGTGGGAGTACCCGGTGTGGGGACGATATGTCTGACCCGGACAAAGTCTTA

1156 NCOI,

GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu
1202 GAAATCACCCCTGACGCACCCAGTCACCAATACATCATGACATGCGTTCGGGCGACCTG
CTTTAGTGGGACTCGGTGGGTACGTGGTATTATGTAGTACTGTACGTACAGCCGGCTGGAC

1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

GluValValThrS rThrTrpValL uValGlyGlyValLeuAlaAlaLeuAlaAlaTyr
1262 GAGGTCTGCAGCAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCGCGGTAT
CTCCAGCAGTGTCTGTGGACCCAGCAGCAACCCCGCAGGACCGACGAAACCGGCGCATA

1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysPr Ala
 TGCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGGAAGCCGGCA
 ACGGACAGTTGTCCGACGCCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCCT
 1375 NAEI,
 1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln
 ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCCTCTCAG
 TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTC
 1391 DRD1,
 1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu
 CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC
 GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAG
 1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn
 GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAAC
 CCGGAGGACGCTGTGGCGAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG
 1508 PSTI, 1513 TTH3I,
 1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln
 TGGCAAAAACCTCGAGACCTTCTGGCGGAAGCATATGTGGAATTCATCAGTGGGATACAA
 ACCGTTTTTGTAGCTCTGGAAGACCGCTTCGTATACACCTTGAAGTAGTCACCTATGTT
 1571 XHOI, 1592 NDEI,
 1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe
 TACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT
 ATGAACCGCCGAACAGTTGCGACGGACCATTTGGGCGGGTAACGAAGTAATACCGAAAA
 1649 BSTE2,
 1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly
 ACAGCTGCTGTCAACAGCCCACTAACCACTAGCCAAACCCCTCCTCTTCAACATATTGGGG
 TGTGACGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC
 1683 ALWN1 PVU2,
 1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly
 GGGTGGGTGGCTGCCAGCTCGCCGCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC
 CCCACCCACCGACGGGTCGAGCGGCGGGGCCACGGCGATGACGGAACACCCGCGACCG
 1800 ESI1,
 1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla
 TTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA
 AATGACCGCGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT
 1808 KAS1 NARI,
 1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal
 GGGTATGGCGCGGCGCTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC
 CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTGCCCATCCAG

1884 SACI, 1905 BSPH1,

1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal
CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTA
GGAGGTGCCTCTCGGACCACTTAGATGACGGGCGGTAGGAGACGGGGCCCTCGGGAGCAT

1934 TTH3I,

1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal
GTCGGCTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTG
CAGCCGACACAGACAGCTCGTTATGACGCGGCGGTGCAACCGGGGCGGCTCCCCCGTCA

2010 NAEI, 2023 SMAI XMAI,

2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis
CAGTGGATGAACCGGCTGATAGCCTTCGCCCTCCCGGGGAACCATGTTTCCCCCAGCCAC
GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCTTGGTACAAAGGGGGTGGCTG

2073 SMAI XMAI, 2099 DRA3,

2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal
TACGTGCCGGAGAGCGATGCAGCTGCCCGCTCACTGCCATACCTACGACGCTCACTGTA
ATGACAGGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACAT

2121 PVU2,

2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer
ACCCAGCTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCACTCCATGCTCC
TGGGTCGAGGACTCCGCTGACGTGGTCACTATTTCGAGCCTCACATGGTGAGGTACGAGG

2165 ALWN1, 2170 MST2,

2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr
GGTTCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC
CCAAGGACCGATTCCCTGTAGACCTGACCTATACGCTCCACACTCGCTGAAATTCCTGG

2226 ECON1,

2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg
TGGCTAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCTTTGTGTCTGCCAGCGC
ACCGATTTTCGATTGAGTACGGTGTGACGGACCTAGGGGAAACACAGGACGGTGC

2291 ESP1, 2306 PVU2, 2316 BAMHI,

2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMethHisThrArgCysHisCysGlyAla
GGGTATAAGGGGGTCTGGCGAGGGGACGGCATATGCACACTCGCTGCCACTGTGGAGCT
CCCATATTTCCCCAGACCGCTCCCTGCGGTAGTACGTGTGAGCGACGGTGACACCTCGA

2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg
GAGTCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGG
CTCTAGTGACCTGTACAGTTTTTCCTGCTACTCTAGCAGCAGGATCCTGGACGTCC

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

2462 AsnMetTrpSerGlyThrPhePr IleAsnAlaTyrThrThrGlyProCysThrProLeu
AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACACGGGGCCCTGTACCCCCCTT
TTGTACACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGAA

2486 ASE1, 2503 APA1,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle
 CCTGCGCGGAACACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA
 GGACGCGGCTTGATGTGCAAGCGGATACCTCCACAGACGTCTCCTTATGCACCTCTAT

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro
 AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCG
 TCCGTCCACCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGG

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe
 TGCCAGGTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCCTACATAGGTTT
 ACGGTCAGGGTAGCGGGCTTAAAAGTGCTTAACTGCCACGCGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu
 CGCGCCCTTCGAAGCCCTTGCTGCGGGAGGAGGTATCATTACAGTAGGACTCCACGAA
 CGCGGGGGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer
 TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCC
 ATGGGCCATCCAGCGTTAATGGAACGCTCGGGCTTGGCTGCACCGGCACAACTGCAGG

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly
 ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGGCGGGCGAAGGTTGGCGAGGGGA
 TACGAGTGACTAGGGAGGGTATATTGTGCTCTCCGCGCGCCGCTTCCAACCGCTCCCTT

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla
 TCACCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA
 AGTGGGGGAGACACCGGTGAGGAGCGGATCGGTGATAGGCGAGGTAGAGAGTTCCGT

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp
 ACTTGACCCGTAAACATGACTCCCTGATGCTGATGCTCAGAAAACAAAGTGGTGATTCTG
 TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu
 AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG
 TCCCTCTCTACCCGCGGTTGTAGTGGTCCAACTAGCTCTTTGTTTCCACCTAAGAC

3062 AspSerPheAspProL uValAlaGluGluAspGluArgGluIleSerValProAlaGlu
 GACTCCTTCGATCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGAGAA
 CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTTAGAGGCATGGGCGTCTT

3102 BGL2,

IleLeuArgLysSerArgArgPheAlaGlnAlaLeuPr ValTrpAlaArgPr AspTyr
 3122 ATCCTGCGGAAAGTCTCGGAGATTCCGCCAGGCCCTGCCCGTTTGGGCGCGCCGGACTAT
 TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCTGATA
 3149 ALWN1, 3170 EAG1 XMA3,
 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly
 3182 AACCCCCCGCTAGTGAGACGTGGAAAAGCCCGACTACGAAACCACTGTGGTCCATGGC
 TTGGGGGGCGATCAGCTCTGCACCTTTTTCGGGCTGATGCTTGGTGAGACCAAGTACCG
 3223 HGIE2, 3235 NCOI,
 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal
 3242 TGCCCCGCTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCTCGGAAGAAGCGGACGGTG
 ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCTGCCAC
 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly
 3302 GTCCTCACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAAGAGCTTTGGC
 CAGGAGTGACTTAGITGGGATAGATGACGGAAACCGGCTCGAGCGGTGGTCTTCGAAACCG
 3338 SACI, 3352 HIND3,
 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro
 3362 AGCTCCTCAACTTCGGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTT
 TCAGGAGTTGAAGGCCCTAATGCCCGCTGTTATGCTGTTGATAGGAGACTCGGGCGGGGA
 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly
 3422 TCTGGCTGCCCGCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCTGGAGGGG
 AGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC
 3443 EAM11051,
 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn
 3482 GAGCCTGGGGATCCGGATCTTAGCGACGGGTATGGTCAACGGTCAGTAGTGAGGCCAAC
 CTCGGACCCCTAGGCCCTAGAATCGCTGCCAGTACCAAGTGCACGTATCACTCCGGTTG
 3490 BAMHI, 3491 BSAB1, 3493 BSPE1,
 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro
 3542 CGCGAGGATGTCGTGCTGCTCAATGTCTTACTCTTGAGACAGGCGCACTCGTCACCCCG
 CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACTGTCCGCGTGAGCAGTGGGGC
 3595 DRA3,
 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis
 3602 TGCGCGCGGAAGAACAACGAACTGCCCATCAATGCACTAAGCAACTCGTTGTAGTGCAC
 ACGCGCGCCTCTTGTGCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTG
 3606 SAC2, 3617 ALWN1, 3661 PFLM1,
 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr
 3662 CACAATTTGGTGTATTCCACCACCTCACGCAGTGTCTTGCCAAAGGCAGAAAGGTACAC
 GTGTTAAACCACATAAGGTGGTGGAGTGCCTCAGCAACGGTTTCCGTCTTCTTCAGTGT
 3687 DRA3,
 PheAspArgL uGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA
AAACTGTCTGACGTTCAAGACCTGTCCGGTAATGGTCTGCATGAGTTCCCTCCAATTTCTGT

AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro
3782 GCGGCGCTCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCGCTGACGCCC
GCGCGAGTTTTCACCTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGG
3822 HIND3,

ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg
3842 CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACCTTCCGTTGCCATGCCAGA
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCT
3881 AAT2, 3896 BGLI,

LysLysAlaThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro
3902 AAGGCGGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly
3962 ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCGTAGAAGGGGGT
TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCA

ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet
4022 CGTAAGCCAGCTCGTCTCATCGTGTCCCGATCTGGGCGTGC GCGTGTGCGAAAAGATG
GCATTTCGGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACGCTTTTCTAC

AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe
4082 GCTTTGTACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTG
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACCTACCCCTCGAGGATGCCTAAG

GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr
4142 CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC
GTTATGAGTGGTCTGTCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGG
4166 ECORI,

ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle
4202 CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC
GGTTACCCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAG
4235 DRD1, 4242 ALWN1,

ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle
4262 CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCGCGTGCCATC
GCATGCCCTCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAG
4307 BGLI, 4314 BALI,

LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn
4322 AAGTCCCTCACCGAGAGGCTTTATGTTGGGGGCCCTTACCAATTCAAGGGGGGAGAAC
TTCAGGGAGTGCTCTCCGAAATACAACCCCCGGGAGAATGTTTAAAGTCCCCCTCTTG
4351 APAI,

CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu
4382 TGC GCGTATCGCAGGTGCCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCTC

ACGCCGATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTGTGGGAG

4442 ThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet
 ACTTGCTACATCAAGGCCCGGGCAGCCTGTCTGAGCCGCGAGGGCTCCAGGACTGCACCATTG
 TGAACGATGTAGTTCGGGGCCCGTCTGGACAGCTCGGCGTCCCGAGGTCTCTGACGTGGTAC
 4458 SMAI XMAI,

4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla
 CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGCTCCAGGAGGACGCG
 GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGC
 4514 DRD1, 4517 TTH3I,

4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro
 CGCAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCG
 CGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGCGGGGGGGACCCCTGGGG
 4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla
 CCACAACCAGAAATACGACTTGGAGCTCATAACATCATGCTCTCCACAGTGTGAGTCGCC
 GGTGTTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGTTGCACAGTCAGCGG
 4643 SACI,

4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla
 CACGACGGCGCTGGAAGAGGGTCTACTACCTCACCGTGACCCCTACACCCCCCTCGCG
 GTGCTGCCGCGACCTTTCTCCAGATGATGGAGTGGGCCTGGGATGTTGGGGGGAGCGC
 4737 NRUI,

4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle
 AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTGTGGTAGGCAACATAATC
 TCTCGACGCACCTCTGTCGTTCTGTGTGAGGTCAAGTAAAGACCGATCCGTTGTATTAG
 4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu
 ATGTTTGGCCCCACACTGTGGGCGAGGATGATACTGATGACCACTTTCTTTAGCGTCCTT
 TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAA
 4812 PFLM1, 4813 DRA3,

4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer
 ATAGCCAGGGACCACTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC
 TATCGGTCCCTGTCGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG
 4899 BGL2,

4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer
 ATAGAACCCTGGATCTACCTCCAATCATTCAAAGACTCCATGGGCTCAGCGCATTTTCA
 TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCCGCTAAAGT
 4960 NCOI,

4982 LeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly
 CTCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGATGCTCAGAAAATTGGG
 GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCACCGCGTACGGAGTCTTTTGAAACC
 5021 SPHI, 5041 KPNI,

5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla
 GTACCGCCCTTGCGAGCTTGAGACACCGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC
 CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGCCCTCGCAGGCGCGATCCGAAGACCGG
 5070 APAI, 5097 BALI,
 ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys
 5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAAGACAAAG
 TCTCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTC
 5119 NDEI,
 LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla
 5162 CTCAACTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT
 GAGTTTGAGTGAGTTTATCGCCGGCGACCGTCGACCTGAACAGGCCGACCAAGTGCCGA
 5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,
 GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp
 5222 GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGG
 CCGATGTCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACC
 5246 DRA3,
 PheCysLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
 5282 TTTTGCTACTCCTTGCTGTCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAAGG
 AAAACGGATGAGGACGAACGACGTCCTCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC
 5301 PSTI, 5331 HGIE2,
 5342 TTGGGGTAAACACTCCGGCCTAAAAAATACTAGAACCCGAGTCGAC
 AACCCCATTTGTGAGGCCGATTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG
 5378 XBAI, 5390 SALI,

FIG. 11-Page 9

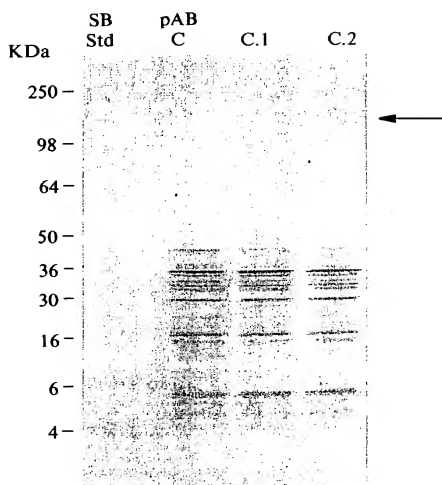


FIG. 12

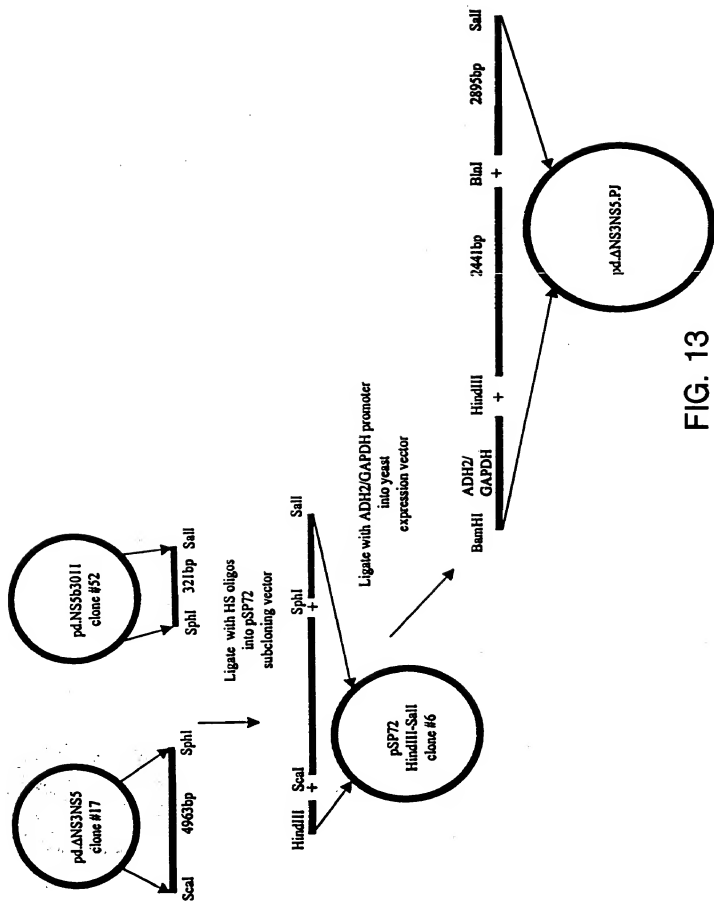


FIG. 13

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTGTGTTTACCGACGTATACGTTCGAGTCCCGATATTCCACGATCATGAGTTG
 ^
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCTCTGTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGGACCAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA
 ^
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCCTGGCAGCCCCATCAGTACTCCACC
 GGATTGTAGTCTCGGCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCTTGGCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAAATAATTGT
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGGCACTCCACGGATGCCACATCCATCTTGGGCATTGGCTGCTCTTGACCAA
 CTGCTACCGGTGAGGTGCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACCTGGT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCCGGGGCGAGACTGGTTGTGCTCGCCACC GCCACCCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCATCCCAACATCGAGGAGTTGCTCTGTCCACCACGGAGAGATCCCTTTT
 TGACACGGGTAGGTTGTAGTCTCTCCAACGAGACAGGTGGTGGCCTCTCTAGGAAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTCTGTCTAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCCTCTGTAGAGTAGAAGACAGTA
 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGATTGGGCATCAATGCCGTG
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGCTCTTGACGTGTCGGTCAACCGACCGGCGATGTTGTCTGCTGCT
 CGGATGATGGCGCCAGAATGCGACAGGCAGTAGGGCTGGTTCGCGCTACAACAGCAGCAC
 ^
 550 SAC2, 560 DRD1,
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCCGTTGATAGACTGCAAT
 CGTTGGCTACGGGACTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
 ^
 615 BSPH1,
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

662 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCCTACCTTCACCAATTGAGACAATC
TGCACACAGTGGGTCTGTACGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACCTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
ACAGTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
TGCAGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTTCGACTCGTCC
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCGCCGAGACT
CAGGAGACACTCACGATACTGCGTCCGACACGAACATACTCGAGTGCGGGCGGCTCTGA

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGCTTCCCGTGTGCCAGGACCATCTT
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTGGTAGAA

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCAG
CTTAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAGATAGGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGGCGT
TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGTTTCGGTGGCACACGCGA

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
AGGGCTCAAGCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTTCGCCCTCAAG
TCCCGAGTTCGGGGAGGGGGTAGCACCTTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
CCCACCTCCATGGGGCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTGAATC
GGGTGGGAGGTACCCGGTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
TGGGACTCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
GTCACGAGCACCTGGGTGCTCGTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGGCTG
CAGTGCTCGTGGACCCACGAGCAACCGCCGACGACCGCAACCGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAAGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCCAGTAGTATCACCCGTCCACGAGAACAGGCCCTTCGGCCGCTTAGTAT
 1369 NAEI,
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTAGAGGGAAGTCCCTCTACCGAGATTTCGATGAGATGGGAAGTGCCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCTGTAAT
 1385 DRD1,
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCTGCTTCCGGGAGCCGGAG
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 1502 PSTI, 1507 TTH3I,
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
 1565 XHOI, 1586 NDEI,
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCAATTGGGGCGGTAACGAAGTAACACCGAAATGTCTGA
 1643 BSTE2, 1677 ALWN1 PVU2,
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTCAACAGCCCACTAACCACTAGCCAAACCCCTCCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTGGGGTATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCAGC
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTGAGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 1794 ESPI,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGGCCGCCCATCGGCAGTGTGGACTGGGAAGGTCTCATAGACATCCTTGCAAGGGTAT
 CCGCGCGGTAGCCGTACAACCTGACCCCTTCAGGAGTATCTGTAGGAACGTCCCATTA
 1802 KASI NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGGCGGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 1878 SACI, 1899 BSPH1,

ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
 TGCCTCCTGGACCACTAGATGACGGGGCGTAGGAGAGCGGGCCCTCGGGAGCATCAGCCG
 1928 TTH3I,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATCTACTGCGCGGGCACGTTGGCCCGGGCGAGGGGGCAGTGACAGTGG
 CACCAGACACSTCGTTATGACGCGGGCGTGCAACCGGGCCCTCCCCGTCACGTCACC
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATCTCAGCAGCCTCACTGTAACCCAG
 GGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGGAGTGACATTGGGTC
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCTCGAGCGCACTGCACCGTGGATAAGCTCGGAGTGACCACTCCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTCACTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGCACTTTAAGACCTGGCTA
 ACCGATTCCCTGTAGACCTGACCTACGCTCCACAACCTCGTGAACCTTGGACCGAT
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCTTTGTGTCTGCCAGCGCGGGTAT
 TTTCGATTTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT
 2285 ESP1, 2300 PVU2, 2310 BAMHI,
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCGACACCGCTCCCTCGCGTAGTACGCTGTGAGCGGAGCTGACACCTCGACTCTAG
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCCTGCTACTCTAGCAGCCAGGATCTGGACGCTCTTGATC
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProL uProAla
 2462 TGGAGTGGGACCTTCCCCATTAAATGCCTACACCAGGGGCCCTGTACCCCCCTTCTCGG
 ACCTCACCTCGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC
 2480 ASE1, 2497 APAI,
 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
 2553 PSTI,
 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAACTCTTAAATGCCCCGTGCCAG
 CACCCCTGAAGGTGATGCACCTGCCCATACTGATGACTGTTAGAATTACGGGGACGGTC
 2594 DRA3,
 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 2642 GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCCTACATAGGTTTGCGCCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG
 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 2702 CCTGCAAGCCCTTGCTCGGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
 GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC
 2757 HGIE2,
 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 2762 GTAGGGTCGCAATTACCTTGCGAGGCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
 CATCCAGCGGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG
 2809 AAT2,
 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 2822 ACTGATCCCTCCCATATAACAGCAGAGGCGGGCGGGCGAAGGTTGGCGAGGGGATCACCC
 TGA CTAGGGAGGGTATATTGTCGTCTCCGCGGGCCGCTTCCAACCGCTCCCTAGTGGG
 2850 EAG1 XMA3,
 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 2882 CCCTCTGTGGCGAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTCCGTTGAACG
 2889 BALI, 2903 NHEI,
 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 2942 ACCGCTAACCATGACTCCCCCTGATGCTGAGCTCATAGAGGCCAAGCTCCTATGGAGGCAG
 TGGCGATTGTTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
 2966 ESPI, 2969 SACI,
 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCGCGCTTGATGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 3062 TTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
 AAGCTAGGCGAACACCGCCTCCTCTGCTCGCCCTCTAGAGGCATGGCGCTTTTAGGAC
 3096 BGL2,
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 3122 CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCGTTTGGGCGCGCGGCACTATAACCCC

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCCGCGGCTGATATTGGGG
 3143 ALWN1, 3164 EAG1 XMA3,
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACACCTGTGGTCCATGGCTGCCCG
 GGCGATACCTCTGCACCTTTTTCGGGCTGATGTTTGGTGACACCAAGTACCGACGGGC
 3217 HGIE2, 3229 NCOI,
 LeuProProLysSerProProValProProProArgLysLysArgThrValValLeu
 3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCTCTC
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAAGGAG
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 3302 ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAAGCTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
 3332 SACI, 3346 HIND3,
 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 3362 TCAACTTCGGGATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGGC
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 3422 TGCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCCTGGAGGGGGAGCCT
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
 3437 EAM11051,
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 3482 GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCACTAGTGAGGCCAACCGGGAG
 CCCCTAGGCGCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCACTCCGGTTGCGCCTC
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTGTCACCCCGTGCGCC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGGCG
 3589 DRA3, 3600 SAC2,
 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
 3602 GCGGAAGAAGACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAAT
 CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTGGTGTTA
 3611 ALWN1, 3655 PFLM1,
 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
 3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCACATTTGAC
 AACCATAAAGGTGGTGGAGTGCCTCACGACGGTTTCCGCTCTTCTTCAGTGTAACCTG
 3681 DRA3,
 ArgLeuGlnValLeuAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
 3722 AGACTCAAGTCTTGACAGCCATTACCAAGGAGCTACTCAAGGAGGTTAAAGCAGCGCGC
 TCTGACGTTCAAGACCTGTCCGTAATGGTCTCGATGAGTTCCTCCAATTTCTGCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
 TCAAAAGTGAAGGCTAACTTGTCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC
 AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGGGGTGTG
 3816 HIND3,
 3842 SerAlaSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
 TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGCTCCGTTGCCATGCCAGAAAGGCC
 AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTCGAGGCAACGGTACGGCTTTCCGG
 3875 AAT2, 3890 BGLI,
 3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
 GTAACCCACATCAACTCCGTGTGGAAGACCTTCTGGAAGACAATGTAACCAATAGAC
 CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG
 3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
 ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTGAGCCTGAGAAGGGGGTCTGTAAG
 TGATGGTAGTACCGATTCTTGTCTCAAAGACGCAAGTCGGACTCTTCCCCCAGGATTC
 4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
 CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCAGCTGTGCGAAAGATGGCTTTG
 GGTGAGCAGAGTAGCAAGGGGCTAGACCCGCACGCGACAGCGCTTTTCTACCGAAAC
 4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
 TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
 ATGCTGCACCAATGTTTCGAGGGGAACGGGCACTACCCTTCGAGGATGCCTAAGTTATG
 4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
 TCACCAAGGACAGCGGGTTGAATTCTCGTGCAAGCGTGGAAGTCCAAGAAAAACCAATG
 AGTGGTCTGTGCGCCAACTTAAGGAGCACGTTTCGACCTTCAGGTTCTTTGGGGTTAC
 4160 ECORI,
 4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
 GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
 CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGAGGCATGC
 4229 DRD1, 4236 ALWN1,
 4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATCAAGTCC
 CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGCGGCGCACCGGTAGTTCAGG
 4301 BGLI, 4308 BALI,
 4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
 CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTCGCGC
 GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCTCCCTCTTGACGCCG
 4345 APAI,
 4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
 TATCGAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCTCACTTGC
 ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTTGTTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
 4442 TACATCAAGGCCCGGGGACGCTGTCTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG
 ATGTAGTTCCGGGCCGCTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC
 ^
 4452 SMAI KMAI,
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
 4502 TGTGGCGACGACTTAGTCGTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
 ACACCGCTGCTGAATCAGCAATAGACACTTTTCGCGCCCCAGGTCTCTCTGCGCCGCTCG
 ^ ^
 4508 DRD1, 4511 TTH3I,
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA
 GACTCTCGGAAGTGCTCCGATCTGGTCCATGAGCGGGGGGACCCCTGGGGGGTGTT
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
 4622 CGAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCACGAC
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
 ^
 4637 SACI,
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
 4682 GCGCGTGGAAAGAGGGTCTACTACCTCACCCGTGACCCACAAACCCCTCGCGAGAGCT
 CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTGGGGGGAGCGCTCTCGA
 ^
 4731 NRUI,
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
 4742 GCGTGGGAGACAGCAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATCATGTTT
 CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
 4802 GCCCCGACACTGTGGGCGAGGATGATACTGATGACCCATTCTTTAGCGTCCTTATAGCC
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAGAAATCGCAGGAATATCGG
 ^ ^
 4806 PFLM1, 4807 DRA3,
 ArgAspGlnLeuGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
 4862 AGGGACCCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA
 TCCCTGGTGAAGTTGTCCGGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT
 ^
 4893 BGL2,
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
 4922 CCACTGGATCTACTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
 GGTGACCTAGATGGAGGTAGTAAGTTTCTGAGGTACCGAGTCGCGTAAAGTGAGGTG
 ^
 4954 NCOI,
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGATGCCTCAGAAAACTTGGGGTACCG
 TCAATGAGAGGTCCACTTTAGTTATCCACCGCGCTACGGAGTCTTTTGAACCCCATGGC
 ^
 5015 SPHI, 5035 KPNI,
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

5042 CCCTTGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
 GGGAACGCTCGAACCTCTGTGGCCCGGCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
 5064 APAI, 5091 BALI,
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAGCTCAAA
 CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCGCTCATTCTTGTTTCGAGTTT
 5113 NDEI,
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
 5162 CTCACCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACAGGCTGGCTAC
 GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
 5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
 5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC
 TCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAACG
 5240 DRA3,
 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
 5282 CTACTCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCAGTGAATAGTCGAC
 GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTATCAGCTG
 5295 PSTI, 5336 SALI,

FIG. 14-Page 9

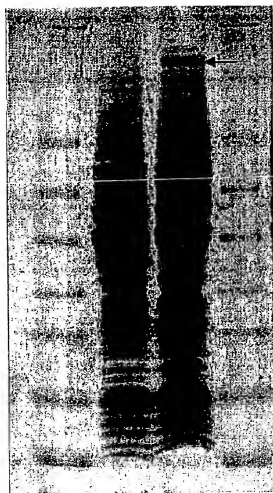
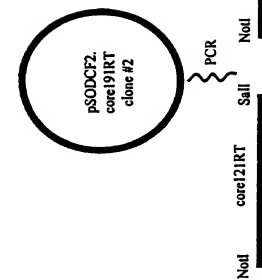
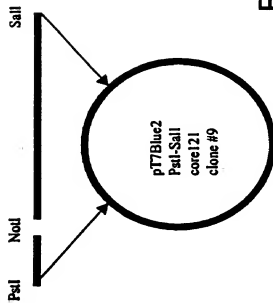


FIG. 15



Ligate with PN oligos
into pT7Blue2
subcloning vector



Ligate with PN oligos
into pT7Blue2
subcloning vector

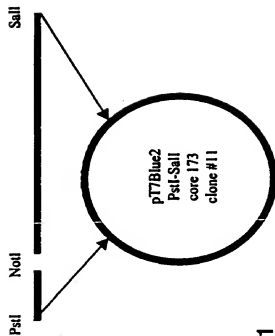
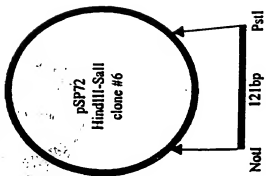


FIG. 16-Page 1



Ligate fragments into pd.ΔNS3NS5.PJ
NotI-Sall cloning vector.



FIG. 16-Page 2

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTTGTTTACCGACGTATACGTCGAGTCCCGATATCCACGATCATGAGTTG
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCTCTGTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC
 GGATTGTAGTCCTGCCCCACTCTTGTAAATGGTGACCGTCGGGGTAGTGATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCCTTCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGGCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
 CTGCTCACGGTGAGGTGCCTACGGGTGAGGTAGAACCCTAACCGTGACAGGAACCTGGTT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCCATCCCAACATCGAGGAGTTGCTCTGTCCACCCGAGAGATCCCTTTT
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTCTGTGCAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCTCCCTCTGTAGAGTAGAAGACAGTA

FIG. 17-Page 1

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAGAAGAAGTGGCAGCAACTCGCCGCAAAGCTGGTTCGCATTGGGCATCAATGCCGTG
 AGTTTCTTCTTACCGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTGTCGTG
 CGGATGATGGGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGTACAACAGCAGCAC

550 SAC2, 560 DRD1,

AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGCTATACCGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA

615 BSPH1,

ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTCAACCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATGAGACAATC
 TGCACACAGTGGGTCTGTAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACTCTGTAG

ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGCGAGGACTGGCAGGGGGAAG
 TGCGAGGGGGTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTCCGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGCCCCCTCGCGGGGAGGGCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCTCTGTGAGTGTATGACGAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
 CAGGAGACACTCACGATCTGCGTCCGACACGAACCTACTCGAGTGCGGGCGGCTCTGA

881 SACI,

ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCAATCCGATGCTCGCATGTACTGTGGGGGCCCGAAGGGCACACGGTCTGGTAGAA

931 SMAI XMAI,

GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTCTATCCGAG
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAGATAGGGTC

985 STUI,

ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAGCCACCGTGTGCGCT
 TGTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

ArgAlaGlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 CCCACCTCCATGGGCCAACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
 ^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGTCATGTGGGCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
 ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 GTCACGAGCACCTGGGTGCTCGTTGGCGCGCTCCTGGCTGCTTTGGCCGCGTATTGGCTG
 CAGTGTCTGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGCGCATACCGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 TCACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGCAATCAT
 AGTTGTCCGACGCACAGTATCACCCGTCCACGAGAACAGGCCCTTCGGCCGTAGTAT
 ^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTCTCTCAGCACTTA
 GGATGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT
 ^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCCTC
 GGCATGTAGTCTGTTCCCTACTACGAGCGGCTCGTCAAGTTGCTCTTCCGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA
 GAGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 AACTCTGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCTATGTTATGAAC
 ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 GCGGGCTTGTCACCGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCGAACAGTTGCGACGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA
 ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 GGTGTCACGAGCCCACTAACCACTAGCCAAACCTCCTCTTCAACATATGGGGGGGTGG
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 GTGGCTGCCAGCTCGCCGCCCGGTGCCGTACTGCCTTTGTGGGCGTGGCTTAGCT
 CACCGACGGGTGCGACGGCGGGGGCCACGGCGATGACGGAACACCCGCGACCGGAATCGA
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGGCGCGCCATCGGCAGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGCAAGGTAT
 CCGCGCGGTAGCCGTCACAACCTGACCCCTTCAGGAGTATCTGTAGGAACGTCCCAT
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGC CGCGGGCGTGGCGGAGCTCTTGTGGCATTCAAGATCATGACGGTGAGGTCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACCCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 1878 SACI, 1899 BSPH1,
 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGAGGCCCTCGTAGTCGGC
 TGCCTCTGGACCACTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
 1928 TTH3I,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCAGTTGGCCCGGGCAGGGGGCAGTCAGTGG
 CACCAGACACGTCGTATGACGCGGCCGTGCAACCGGGCCCGTCCCCCGTCACGTCAAC
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
 TACTTGGCCGCTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGCGAGCTGCCCGCGTCACTGCCATACCTACGACGCTCACTGTAACCCAG
 GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCGCCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTGCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
 ACCGATTCCTGTAGACCTGACCTATACGCTCCACAACCTCGCTGAAATTCGGACCGAT
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGGGGTAT
 TTTGATTGAGTACGGTGTGACGGACCTAGGGGAAACACAGGACGGTCCGCGCCATA
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCTGTACTCTTAGCAGCCAGGATCTGTGACGTCCTTGTAC

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 TGGAGTGGGACCTTCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTTCTCGCG
 ACCTCACCTTGAAGGGTAAATTACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 CGAACTACACGTTCCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
 CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 GTCCCATCGCCCAATTTTTTACAGAATTGGACGGGGTGCCTACATAGGTTTGCGCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCACCGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 CCTGCAAGCCCTTGTGTCGGGAGGAGGTATCATTACAGATAGGACTCCACGAATACCCG
 GGGACGTTCCGGGAACGACGCCCTCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 GTAGGGTCGCAATTACCTTGCGAGCCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCTGCACCGGCACAACCTGCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 ACTGATCCCTCCCATATAACAGCAGAGGGCGGCCGGGGGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGAGGGTATTATTGTCGTCTCCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProS rLeuLysAlaThrCys
 CCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTCTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCGGTTGAACG

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
 2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCCGCGTTGTAGTGGTCCCACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 TTCGATCCGCTTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
 AAGCTAGGCGAACACCGCCTCCTCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
 3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCGCTTTGGGCGCGGCCGACTATAACCC
 GCCTTACAGAGCCTCTAAGCGGGTCCGGGACGGGCAAAACCCGCGCGCTGATATTGGGG

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGGTGCCCG
 GGCATCACCTCTGCACCTTTTTCGGGCTGATGCTTTGGTGGACACAGGATACCGACGGGC

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 CTTCACCTCCAAAGTCCCCTCCTGTGCCTCCGCTCGGAAGAAGCGGACGGTGGTCTCT
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTCGCCTGCCACGAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 ACTGAATCAACCTTATCTACTGCCTTGGCCGAGCTCGCCACCAGAGGTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 TCAACTTCGGGATTTAGCGGCGCAACATACGACAACACTCCTAGAGCCGCCCTTCTGGC
 AGTTGAAGGCCGTAATGCCGCTGTTATGCTGTTGTAGGAGACTCGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 TGCCCCCGGACTCCGACGCTGAGTCTATTCTCTCCATGCCCCCGCTGGAGGGGGAGGCT
 ACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 GGGGATCCGGATCTTAGCGAGCGGTCTATGGTCAACGGTCAAGTAGTGAGGCCAACCGGAG
 CCCCTAGGGCTAGAATCGTGCCAGTACCAGTTGCCAGTCATCACTCCGTTGCGCCTC

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCAACCCGTGCGGCC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGACAGTGGGGCAGCGGC

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAA
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTAGCAACGATGCAGTGGTTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGTATTCCACCACTCACGCAGTGCTTGCCAAAGGCAGAAAGTCACATTGAC
AACCACATAAGGTGGTGAGTGCGTCACGAACGGTTTCCGTCTTCTTCAGTGTAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCTCCAAATTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAGTGAAGGCTAAGTTCGCTATCCGTAGAGGAAGCTTCGACGCTGACGCCCCAC
AGTTTTCTACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGGAAAAGACGTCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCGTTCACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTCTGCGTTACGCTGAGAAGGGGGGTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGATGACACAGGGGCTAGACCCGCACGCGCACAGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAGACTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACAGGACAGCGGGTGAATTCTCGTGCAAGCTGGAAGTCCAAGAAACCCCAATG
AGTGGTCTGTGCGCCAACTTAAGGAGCAGCTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTCGCGGCACCCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGC
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCTCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
TATCGCAGGTGCCGCGCAGCGCGCTACTGACAACTAGCTGTGGTAACACCCCTCACTTGC
ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
TACATCAAGCCCGGGCAGCCTGTGAGCCGCGAGGGTCCAGGACTGCACCATGCTCGTG
ATGTAGTTCGCGGCCGCTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGAGC
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCACAA
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
CCGAATACGACTGGAGCTCATAACATCATGCTCCTCCAAAGTGTGTCAGTCCGCCACGAC
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTGACACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACACCCCCCTCGCGAGAGCT
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
GCGTGGGAGACGACAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATCATGTTT
CGCACCTCTGTGCTGTTGTGTGAGGTGAGTTAAGGACCGATCCGTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePh SerValLeuIle Ala
GCCCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTTAGCGTCTTATAGCC
CGGGGGTGTGACACCCGCTCTCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIle TyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCCTGCTACTCCATAGAA
 TCCTTGGTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGAGTCGCGTAAAGTGAGGTG

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGATGCCTCAGAAAACCTTGGGGTACCG
 TCAATGAGAGGTCCACTTTAGTTATCCCACCGCGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
 5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
 GGGAAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCACTGGGCAGTAAGAACAAGCTCAA
 CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
 5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC
 GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCCAAGTGGCCGACCGATG

5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
 5222 AGCGGGGGAGACATTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTC
 TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
 5282 CTACTCTGCTTGTGCGAGGGGTAGGCATCTACCTCTCCCCAACCGAATGAGCAGCAAT
 GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
 5342 CTAACACCTCAAAGAAGACCAACGTAACACCAACCGCGGCCGAGGACGTCAGTTTC
 GGATTTGGAGTTTCTTTCTGTTTGCATTGTGGTTGGCCGCCCGGCTCTGCGAGTTCAAG

5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
 5402 CCGGCTGGCGGTACAGATCGTTGGTGGAGTTTACTTGTTCGCGCGCAGGGGCCCTAGATTG
 GGCCACCGCCAGCTAGCAACCACTCAAATGAACAACGGCCGCTCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTGAGCCT
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGTACCCTTGGCCC
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGCGCAATTTGGGTAAGTAATAGTCG
GGATCGACCCCGGGGTGTCTGGGGGCGCATCCAGCGCGTTAAACCCATTCAATATCAG

5650 APAI, 5698 SALI,

5702 AC
TG

FIG. 17-Page 10

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
TCGAATGTTTTGTTTACCGACGTATACGTCCGAGTCCCGATATCCAGCATGAGTTG
1 HIND3, 24 NDEI, 52 SCAI,
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
62 CCCTCTGTGTGCTGCAACACTGGGCTTTGGTGTCTTACATGTCCAAGGCTCATGGGATCGAT
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCTAGCTA
116 CLAI,
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCAGTACTCCACC
GGATTGTAGTCTTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
182 TACGGCAAGTTCCCTTGC CGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTGT
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAACA
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
CTGCTCAGGTGAGGTGCTACGGTGTAGGTAGAACCCTAACCGTGACAGGAACCTGGTT
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
302 CGCAGAGACTCGGGGGGCGAGACTGGTTGTGCTCGCCACGCCACCCTCCGGGCTCCGTC
CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCGAGGCAG
303 ALWN1,
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
422 TACGGCAAGGTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCTAT
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
482 TCAAAGAAGAAGTCCGACGAACTCGCCGCAAGCTGGTGCATTTGGGCATCAATGCCCGTG
AGTTTCTTCTTACGCTGCTTGAGCGGGGCTTTCGACCAGCGTTAACCCGTAGTTACGGCAC
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
542 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGCGCATGTTGTCTGTCGTG
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTGCCTGCTACAACAGCAGCAC
550 SAC2, 560 DRD1,
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAsp rValIleAspCysAsn
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
CGTTGGCTACGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
615 BSPH1,

ThrCysValThrGlnThrValAspPheS rLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCTACCTTCACCATTGAGACAATC
 TGCACACAGTGGGTCTGTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACCTCTGTTAG

ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 AGCCTCCCCAAGATGCTGTCTCCCGCACTCAACGTGGGGCAGGACTGGCAGGGGGAAG
 TCGAGGGGGTTCACGACAGAGGGCGTGAGTTGACGCCCGCTCCTGACCGTCCCCCTTC

ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTGAGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCTCTGTGAGTGCTATGACGAGGCTGTGCTTGGTATGAGCTCACGCCCGCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCTACTCGAGTGGGGCGGCTCTGA

881 SACI,

ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCGAAGGGCACACGGTCTCGGTAGAA

931 SMAI XMAI,

GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTGGGAGGGCGCTTTACAGGCCCTCACTCATATAGATGCCCATTTCTATCCGAG
 CTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTTCTGTCTACCCCTCTTGAAGGAATGGACCATCGCATGTTCGGTGGCACACGCGA

1069 DRA3,

ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAGCCCCCTCCCCCATCGTGGGACCATGTGGAAGTGTGTTGATTTCGCTCAAG
 TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 1142 CCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTGAGAATGAAATC
 GGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAGTCTTACTTTAG

1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 1202 ACCCTGACGACCCAGTCAACCAATACATCATGACATGCATGTGGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAGCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGCGCTCTGGCTGCTTTGGCCGCGTATTGCTG

CAGTGCTCGTGGACCCACGAGCAACCGCCGACGACGAAACCGGCGCATAACGGAC
 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGAGGGTCTGTTGCCGGAAGCCGGCAATCATA
 AGTTGTCGACGACACAGTATCACCCGTCGACGAGAACAGGCCCTTCGGCCGTTAGTAT
 1369 NAEI,
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
 1385 DRD1,
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCCTGCTCCAGACCACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 1502 PSTI, 1507 TTH3I,
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCCTATGTTATGAAC
 1565 XHOI, 1586 NDEI,
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CCGCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACACGAAAATGTGCA
 1643 BSTE2, 1677 ALWN1 PVU2,
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTCACGAGCCCACTAACCACTAGCCAAACCCCTCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCAGCTCGCCGCCCGGTCGCCCTACTGCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTGAGCGCGGGGGCCACGGCGATACGGAACACCCGCGACCGAATCGA
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGGCCCGCATCGGCAGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGACGGGTAT
 CCGCGCGGTAGCCGTCAACAACCTGACCCTTCCAGGAGTATCTGTAGGAACGTCCATA
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATAGCCGGTGAGGTCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCATCCAGGGGAGG
 1878 SACI, 1899 BSPH1,

ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
 TGCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
 1928 TTH31,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCAGCTTGGCCCGGGCGAGGGGCGAGTGCAGTGG
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCAAC
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCACGCACTACGTG
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGACAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
 GGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTCT
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCACCACTGGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTCACCTATTCCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTGAGCGACTTTAAGACCTGGCTA
 ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACCTCGTGAAATTCTGGACCGAT
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTGTGTCTGCCAGCGCGGGTAT
 TTTCCGATTGAGTACGGTGTGCGACGGACCTAGGGGAACACAGGACGGTGCAGGCCATA
 2285 ESP1, 2300 PVU2, 2310 BAMHI,
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 2342 AAGGGGTTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 2402 ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 2462 TGGAGTGGGACCTTCCCCATTATGCCTACACCACGGGGCCCTGTACCCCCCTTCTCGCG
 ACCTCACCCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
 2480 ASE1, 2497 APAI,

ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCACAGACGCTCCTTATGCACCTCTATTCCGTC
 2553 PSTI,
 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAACTCTTAAATGCCCGTGCCAG
 CACCCCTGAAGTGATGCACTGCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
 2594 DRA3,
 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 2642 GTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG
 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 2702 CCGTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTACAGAGTAGGACTCCACGAATACCGC
 GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC
 2757 HGIE2,
 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 2762 GTAGGGTCGCAATTACCTTGCAGGCCGGAACCGGACGTGGCGGTGTGACGTCCATGTCT
 CATCCAGCGTTAATGGAAACGCTCGGGCTTGCCCTGCACCGGCACAATGCAGGTACGAG
 2809 AAT2,
 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 2822 ACTGATCCCTCCCATATAACAGCAGAGGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGGAGGGTATTATTGTCGTCTCCGCGGGCCGCTTCCAACCGCTCCCTTAGTGGG
 2850 EAG1 XMA3,
 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 2882 CCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTCCGTTGAACG
 2889 BALI, 2903 NHEI,
 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 2942 ACCGCTAACCATGACTCCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
 2966 ESp1, 2969 SACI,
 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 3002 GAGATGGGCGGCACATCACCAGGGTTGAGTCAGAAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
 PheAspProLeuValAlaGluGluAspGluArgGluIleS rValProAlaGluIleLeu
 3062 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAAAATCTGT
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
 3096 BGL2,
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

3122 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAAACCCGCGCCTGATATTGGGG
 3143 ALWN1, 3164 EAG1 XMA3,
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 3182 CCGCTAGTGGGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGCCCC
 GCGCATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGG
 3217 HGIE2, 3229 NCOI,
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 3242 CTTCCACCTCCAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCCTC
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCTGCCACAGGAG
 ThrGluSerThrLeuAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 3302 ACTGAACTCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAACCGTCGAGG
 3332 SACI, 3346 HIND3,
 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 3362 TCACTTCCGGCATTACGGGGGACAAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC
 AGTTGAAGGCCGTAAATGGCCGCTGTTATGCTGTTGAGAGACTCGAGGGGGGAAGACCG
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCTTCTCCTCCATGCCCCCTTGGAGGGGGAGCCT
 ACGGGGGGGCTGAGGCTCGCACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
 3437 EAM11051,
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGGGAG
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTGGCGCCTC
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 3542 GATGTCTGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCAACCCCGTGGCC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCAGCGGG
 3589 DRA3, 3600 SAC2,
 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
 3602 GCGGAAGACAGAACTGCCCATCAATGCACATAAGCAACTCGTTGCTAGCTACCCACAAT
 CGCCTTCTGTCTTTGACGGGTAGTTACGTGATTCTGTTAGCAACGATGCACTGGTGTTA
 3611 ALWN1, 3655 PFLM1,
 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
 3662 TTGGTGTATTCCACCACCTCACGCGAGTGCTTGCCAAAGGCAGAAAGATCACATTGAC
 AACCACATAAGTGGTGGAGTCCGTCACGAACGGTTCCGCTCTTCTTCACTGTAACTG
 3681 DRA3,
 ArgLeuGlnValLys uAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
 3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGGC

TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCTCTCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC
AGTTTTCACCTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCTGTCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCGGTTTCTGACGGCAACGGTACGGTCTTTCCGG
3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGAACACCAATAGAC
TGATGGGTGATGTAGTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTACGCTGAGAAGGGGGTTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAAGGGCTAGACCCGACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGAGCTGGTTTCAAAAGCTCCCTTGCCGTGATGGGAAGCTCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCTCTGTCGAAGCGTGGAAGTCCAAGAAAACCCCAATG
AGTGGTCTCTGCGCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC
4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC
4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATCAAGTCC
CTCCTCCGTTAGTGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG
4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
CTCACCCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG
4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValL uThrThrS rCysGlyAsnThrLeuThrCys
TATCGCAGGTGCCGCGCGAGCGGCTACTGACAACCTAGCTGTGGTAAACACCCTCACTTGC
ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG
ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
TGTGGCGACGACTTGTCTGTTATCTGTGAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCGGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCACAA
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
CCAGAATACGACTTGGAGCTCATACATCATGCTCCCTCAACGTGTGACGTGCGCCACGAC
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
GGCGCTGGAAAGAGGGTCTACTACCTCACCGTGACCTACAACCCCCCTCGCGAGAGCT
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
GCCTGGGAGACAGCAAGACACACTCCAGTCAATTTCCTGGCTAGGCAACATAATCATGTTT
CGCACCTCTGTGTTCTGTGTGAGGTCAAGTAAAGGACCGATCCGTTGATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
GCCCCACACTCTGGGGCGAGGATGATACTGACCCATTCTTTAGCGTCTCTTATAGCC
CGGGGGTGTGACACCCGCTCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA
TCCTTGGTCGAACCTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGAGTCGCGTAAAAGTGAGGTG

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGGCTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
 GGGAAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
 5064 APAI, 5091 BALI,
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAGGCTCAAA
 CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT
 5113 NDEI,
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
 5162 CTCACCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC
 GAGTGAGGTTATCGCGCGGCGGCTCGACCTGAACAGGCGGACCAAGTGCCGACCGATG
 5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 Pvu2,
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
 5222 AGCGGGGGAGACATTATACAGCGTGTCTCATGCCCGGGCCGCTGGATCTGGTTTTCG
 TCGCCCCCTCTGTAAATAGTGTCCGACAGAGTACGGCGCGGGGCGACCTAGACAAAACG
 5240 DRA3,
 LeuLeuLeuLeuAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
 5282 CTACTCCTGCTTGTGTCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCAGCAAT
 GATGAGGACGAACGAGCTCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA
 5295 PSTI,
 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
 5342 CCTAAACCTCAAAGAAAGACCAACGTAACCAACCGGGCGGCGCAGGACGCTCAAGTTC
 GGATTGTGGAGTTCTTTCTGGTTTGATTGTGGTTGGCCCGCGGCTCTGCAGTTCAAG
 5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,
 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
 5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCGCTAGATTG
 GGCCACCGCCAGCTTAGCAACCACCTCAAATGAACAACGGCGCTCCCCGGGATCTAAC
 5449 APAI,
 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCAGAGTAGACGTCAGCCT
 CCACACGCGCGCTGCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
 5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,
 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
 5522 ATCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGATACCTTGGCCC
 TAGGGGTTCCGAGCAGCGGGCTCCCGTCTTGGACCCGAGTCGGGGCCCATGGGAACCGG
 5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,
 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
 5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCTGTCTCCCCGTGGCTCTCGG
 GAGATACCGTTACTCCGACGCCACCGCCCTACCGAGGACAGAGGGGACCGGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
 GGATCGACCCCGGGGTGCTCTGGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA
 5650 APAI, 5696 CLAI,
 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
 5702 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCCTCGTCGGCGCCCTCTT
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA
 5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,
 GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr
 5762 GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTGGAAGACGGCGTGAACATAT
 CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACCTTGATA
 5772 BSTXI, 5775 APAI,
 AlaThrGlyAsnLeuProGlyCysSerOC AM
 5822 GCAACAGGGAACCTTCTGGTTGCTCTTAATAGTCGAC
 CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG
 5854 SALI,

FIG. 18-Page 10

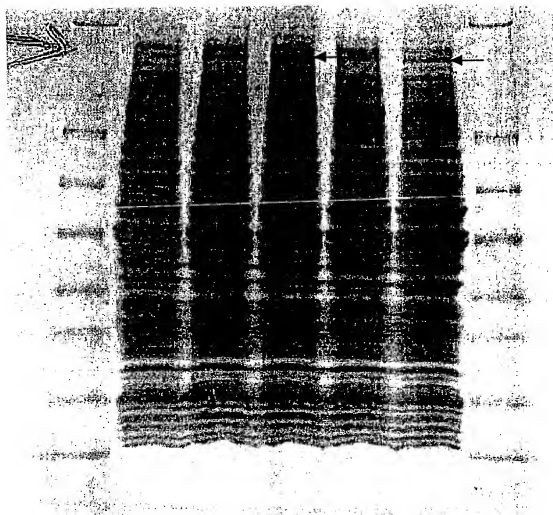


FIG. 19

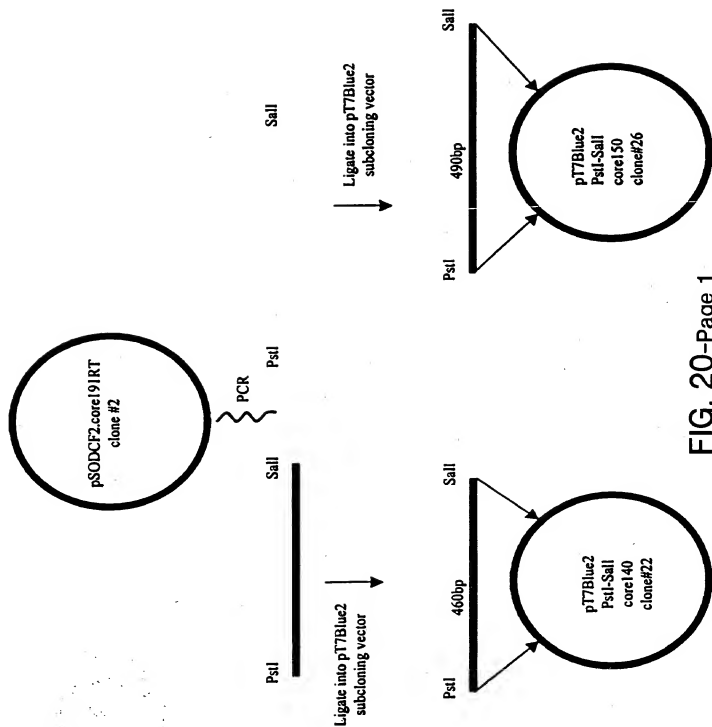
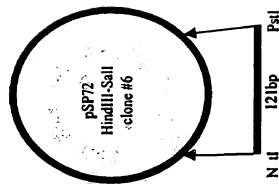


FIG. 20-Page 1



Ligate fragments into pdNS3NS5.PJ
NotI-Sall cloning vector.

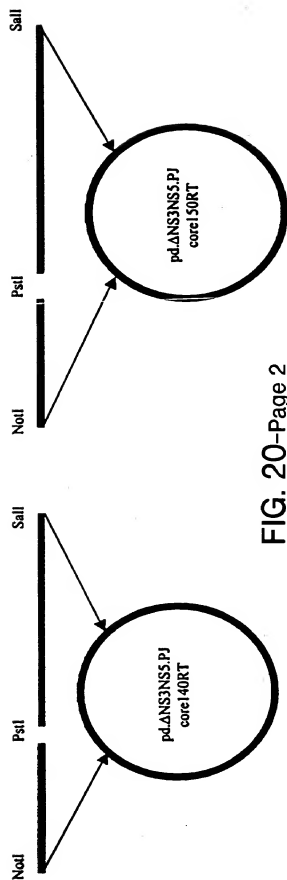


FIG. 20-Page 2

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTGTGTTTACCGACGTATACGTCGAGTCCCGATATTCACGATCATGAGTTG
 ^ ^
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
 ^
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCAGTACTCCACC
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTGTT
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCTTGACCAA
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACC CGTAACCGTGACAGGAAC TGTT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
 ^
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCCATCCCAACATCGAGGAGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
 TGACACGGGGTAGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGAAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTGCAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 21-Page 1

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAAGAAGAAAGTGCACGAACCTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGCTTTGACGTGTCCGTCATCCGACAGCGGCATGTTGTCTGTCGTG
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGTACACAGCAGCAC
 550 SAC2, 560 DRD1,
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGTGAAGCTGAGCCACTATCTGACGTTA
 615 BSPH1,
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTTAGACAAATC
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACTCTGTTAG
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
 TCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGCACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGAGGCCGTACAAGCTGAGCAGG
 816 BGLI, 833 DRD1,
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCCTCTGTGAGTGCTATGACGACGAGCTGTGCTTGGTATGAGCTACGCCCCGCCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATATCTCGAGTGCGGGCGGCTCTGA
 881 SACI,
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGGCACACGGTCTCGGTAGAA
 931 SMAI XMAI,
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTGGGAGGGCGCTTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCAG
 CTTAAACCCCTCCCGCAGAAATGTCGGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
 985 STUI,
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGGCGT
 TGTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
 1069 DRA3,
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGCGCTCAAG

TCCCGAGTTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTGAATC
 GGGTGGGAGGTACCCGGTTGTGGGACGATATGTCTGACCCGCGACAGTCTTACTTTAG
 1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 ACCCTGACGCAACCCAGTCACCAAATACATCATGACATGCATGTGCGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTTCAGTGGTTTATGTAGTACTGTACGTACAGCCGCGCTGGACCTCCAG
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 CTCACGAGGACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCTATTGGCTG
 CAGTGCTCGTGGACCCACGAGCAACCGCGCAGGACCGAGAAACCGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCACAGTATCACCCGTCACAGCAGAACAGGCCCTTCGGCCGTTAGTAT
 1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 CCTGACAGGGAAGTCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGAGAGTCGTGAAT
 1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTTCTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 CTGACAGACCGCTCCCGTCAGGCAGAGGTATCGCCCTGCTGTCACAGCACTGGCAA
 GACGCTCGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCTATGTTATGAAC
 1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 GGGGCTTTGTCACCGCTGCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAAACGAAGTAACTACCGAAATGTCGA
 1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrS rGlnThrLeuL uPheAsnIleLeuGlyGlyTrp
 GCTGTCAACGACCCACTAACCCAGCAACCCCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCAGCTCGCCGCCCGGTCGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGC CGCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
 CGCGCGGGTAGCCGTCACAACTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGC CGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
 CGCGCGCCGACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 1878 SACI, 1899 BSPH1,
 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGGCCCTCGTAGTCGGC
 TGCCTCTGGACAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
 1928 TTH3I,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATCTGCGCGGCACGTTGGCCCGGGCGAGGGGCGAGTGCAGTGG
 CACCAGACAGCTCGTTATGACGCGGGCGTGCAACCGGGCCGCTCCCCCGTCACGTCACC
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCAGCGACTACGTG
 TACTTGGCCGATATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGGGTGATGCAC
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGACGCTGCCCGCTCACTGCCATACCTCAGCAGCCTCACTGTAAACCCAG
 GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGGAGTGACATTGGGTC
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCACAGTGGATAAGCTCGGAGTGTACCCTCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTAAAGACCTGGCTA
 ACCGATTCCTGTAGACCTGACCTATACGCTCCACAACCTCGCTGAAATTCGTGACCGAT
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGGTAT
 TTGCTGATCGATACGGTGTGCGACGGACCTAGGGGAAACACAGGACGGTGCAGGCCATA
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyI1 MetHisThrArgCysHisCysGlyAlaGluIle
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGGAACATG
 TGACCTGTACAGTITTTTGGCCCTGCTACTCTAGCAGCGAGGATCCTGGACGCTCTTGATC

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 ACTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCTTCTCTGCG
 ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 CCGAATCTACAGCTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCACAGCGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGGCCAG
 CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTACGGGCACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 GTCCCATCGCCGAATTTTTCAGAAATTTGGACGGGGTGGCCCTACATAGTTTTGGCCCC
 CAGGGTAGCGGGCTTAAAGAGTGTCTTAACCTGCCCCACGGGATGTATCCAAACGGCGG

2702 CcTcGlyLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 CCTCGCAAGCCCTTGTCTGCGGGAGGAGGTATCATTACAGAGTAGGACTCCACGAATACCCG
 GGGACGTTCCGGAACGACGCCCTCTCCATAGTAGTCTCATCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 GTAGGGTCGCAATTACCTTGCAGGCCGGAACCGGACGTGGCCGTGTTGACGTCCATGCCTC
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAATGCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 ACTGATCCCTCCCATATAACAGCAGAGGCGCGCGGGCGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGAGGGGTATATTGCTGCTCTCCGCGGGCCGCTTCCAACCGCTCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG

2889 BALI, 2903 NHEI,

FIG. 21-Page 5

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC
 2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 GAGATGGGCGGCACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 TTCGATCCGCTTGTGGCGGAGGAGACGAGCGGGAGATCTCCGTACCCGAGAAATCCTG
 AAGCTAGGCGAACACCCGCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
 3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCGTTTGGGCGGGCCGGACTATAACCCC
 GCCTTCAGAGCCTCTCAGCGGGTCCGGACGGGCAAACCCGCCCGGCTGATATTGGGG
 3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
 GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAAGGTACCGACGGCG
 3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 CTTCCACCTCCAAGTCCCTCCTGTGCCTCCGCTCGGAAGAAGCGGACGGTGGTCTCT
 GAAGGTGGAGGTTTCAGGGGAGGACACGAGGCGGAGCCTCTTCGCCTGCCACCGAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
 3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 TCAACTTCGGCATTACGGGCGACAATACGACAACTCCTCTGAGCCCGCCCTTCTGGC
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 TGCCCCCGCATCCGACGCTGAGTCTATTCTCCATGCCCCCTGGAGGGGGAGCCT
 ACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
 3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCAGTAGTGAGCCCAACCGGAG
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAAGTTGCCAGTCATCACTCCGTTGGCGCTC
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrS rTrpThrGlyAlaLeuValThrProCysAla
 GATGTCGTGTGCTCAATGTCTTACTCTTGGACAGCGGCACCTCGTCAACCCGTGCGCC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCTGAGCAGTGGGGCACCGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
CGCCTTCTTGCTTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGATTCCACCACCTCACGCAGTGCCTTGCCAAAGGCAGAGAAAGTCACATTTGAC
AACCCATAAGGTGGTGGAGTGGTCCGTCACGAACGGTTCCGCTCTCTTTCAGTGTAAGCTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCG
TCTGACGTTCAAGACCTGTCCGTAATGGTCTGCATGAGTTCTCCAAATTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC
AGTTTTCACTCCGATTGAACGATAGGCATCTCCTTGAACGTCGGACTGCGGGGGTGTTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACATGTAACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTCATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAAGCAGGTTTTCTCGCTTCAGCCTGAGAAGGGGGGTGCTAAG
TGATGTAGTACCGATTCTTGCTCCAAAGACCAAGTCGGACTGTCGCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTTCGCCGATCTGGGCGTGCGCGTGTGCGAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAGGGGGTAGACCCGCACGCGCACAGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACGAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAGTCCAGAAACCCCAATG
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCCGCTGCTTGACTCCACAGTCACTGAGAGCGCATCCGTACG
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCTGGCCATCAAGTCC
CTCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTT^CGGGCGCACCCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGC
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTC^CCCCCTCTTGACGCCG

4345 AFAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
TATCGCAGGTGCGCGCGAGCGCGTACTGACAACCTAGCTGTGTAACACCTCACTTGC
ATAGCTCCACGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCGAGGGTCCAGGACTGCACCATGCTCGTG
ATGTAGTTC^CGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCTCGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCAGGAGGACGCGCGAGC
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCGCGCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA
GACTCTCGGAAGTGCCCTCCGATACTGGTCCATGAGGCGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCACTGCCCCACGAC
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
GGCGCTGGAAGAGGGTCTACTACCTCACCCTGACCCCTACAACCCCTCGCGAGAGCT
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTTGCTAGGCAACATAATCATGTTT
CGCACCTCTGTCGTTCTGTGTGAGGTCAAGTAAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCTTATAGCC
CGGGGGTGTGACACCCGCTCTACTATGACTACTGGGTAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA
TCCCTGGTCGAACCTTGTCGGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
GGTGACCTAGATGGAGGTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG
TCAATGAGAGGTCCACTTTAGTTATCCCACCGCGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGTAGGCTTCTGGCCAGAGGA
GGGAACGCTCGAACCTCTGTGGCCCGGGCTCGCAGGCGCGATCCGAGACCGGTCTCCT

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAGCTCAA
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACCTCCAATAGCGCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTATCACAGCGTGTCTCATGCCCGGGCCCGCTGGATCTGTTTTGC
TCGCCCCCTCTGTAAATAGTGTCCGACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 TACTCCTGCTTGTCTGCAGGGGTAGGCATCTACCTCTCCCCAACCGAATGAGCACGAAT
GATGAGGACGAACGAGCTCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGCGGCCGAGGACGTCAGATTTC
GGATTGTGAGTTTCTTCTGTTTGCATTGTGGTTGGCCCGCGCTCAGTTCAGT

5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTCGCGCAGGGGGCCCTAGATTG
GGCCACCGCCAGCTAGCAACCACTCAATGAACAACGGCGTCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
 ^ ^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
 ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGGCCCATGGGAACCGGG
 ^ ^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
 GAGATACCGTTACTCCCCAGCGCCACCCGCCCTACCGAGGACAGAGGGGCCACCGAGAGCC
 ^ ^ ^ ^ ^

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
 CCTAGCTGGGGCCCCACAGACCCCGCGCTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
 GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA
 ^ ^ ^ ^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM
 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG
 ^ ^ ^ ^ ^

5724 HGIE2, 5755 SALI,

FIG. 21-Page 10

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
 ^ ^ ^
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
 ^
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGACAATTACCACTGGCAGCCCCATCAGTACTCCACC
 GGATTGTAGTCTGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAAATAATTTGT
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
 CTGCTCACGGTGAGGTGGCTACGGTGTAGGTAGAACCCTAACCGTGACAGGAACCTGGTT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTCGGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCCGTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGGAGGCAG
 ^
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATCCCTTTT
 TGACACGGGGTAGGGTTGTAGTCTCTCCAACGAGACAGGTGGTGGCTCTCTAGGGAAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTGCAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 22-Page 1

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 TCAAAGAAGAGTGGCAGCAACTCGCCGCAAAGCTGGTCGATTGGGCATCAATGCCGTG
 AGTTTCTTCTTACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGCGATGTTGTGTCGTG
 CGGATGATGGGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGTACAAACAGCAGCAC

550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTAGCCACTATCTGACGTTA

615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 ACGTGTGTCAACCGACAGTTCGATTTCAGCCTTGACCCTACCTTCAACATTGAGACAATC
 TGCACACAGTGGGTCTGTGAGCTAAAGTCGGAACCTGGGATGGAAGTGGAATCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTGCGGGCAGGACTGGCAGGGGGAAG
 TGCAGGGGGTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 CCAGGATCTACAGATTTGTGGCACCGGGGGAGCGCCCTCCGGCATGTTCCGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCGCCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGGCGGCTCTGA

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCAATCCGATGTCTCGCATGTACTTGTGGGCCCCGAAGGGCACACGGTCTCTGGTAGAA

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCAG
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 ACAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGATGGTTCCGTGGCACACGCCGA

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 AGGGCTCAAGCCCTCCCCATCGTGGGACCAGATGTGAAGTGTTTGATTGCGCTCAAG

FIG. 22-Page 2

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC
 ProThrLeuHisGlyProThrProL uLeuTyrArgLeuGlyAlaValGlnAsnGluI1
 1142 CCCACCTCCATGGGCCAACACCCTGCTATACAGACTGGGCGCTGTTGAGAATGAAATC
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
 ^
 1150 NCOI,
 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGTCATGTCGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
 ^ ^ ^ ^
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,
 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCGCTG
 CAGTGTCTGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGCGCATAACGGAC
 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCACAGAGGCTGCGTGGTCTAGTGGGCAGGGTCGTCTGTGTCGGGGAAGCCCGCAATCATA
 AGTTGTCGACGCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTAGTAT
 ^
 1369 NAEI,
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA
 GGAAGTCTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT
 ^
 1385 DRD1,
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnAlaLeuGlyLeu
 1442 CCGTAGATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAGGCCCTCGGCCG
 GGCATGTAGTCTGTTCCCTACTACGAGCGGCTCGTCAAGTTCGCTTTCGGGAGCCGGAG
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 ^ ^ ^ ^
 1502 PSTI, 1507 TTH3I,
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAATCTGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAGTAGTCACCTATGTTATGAAC
 ^ ^ ^ ^
 1565 XHOI, 1586 NDEI,
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCTGGTAACCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCGAACAGTTGCGACGACCATTTGGGGCGGTAAACGAAGTAACCTACCGAAAATGTGCA
 ^ ^ ^ ^
 1643 BSTE2, 1677 ALWN1 PVU2,
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPh AsnIleLeuGlyGlyTrp
 1682 GCTGTCAACGACCCACTAACCACTAGCCAAACCCTCCTCTCAACATATTGGGGGGGTGG
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

- 1742 ValAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
GTGGCTGCCAGCTCGCCGCCCGCTGCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
CACCAGCGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
1794 ESP1,
- 1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCTCTATAGACATCCTTGCAGGGTAT
CGCGGGCGGTAGCCGTACACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCTCATA
1802 KAS1 NARI,
- 1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGGGGTGAGGTCCCTCC
CGCGCGCCCGCACCGCCCTCGAGAACCCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
1878 SACI, 1899 BSPH1,
- 1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
TGCTCCTTGAGCACGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
1928 TTH31,
- 1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
GTGGTCTGTGCAGCAATACTGCGCGCGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
CACCAGACACGTCGTTATGACGCGCGCGTGCAACCGGGCCCGCTCCCCCGTCAGCTCACC
2004 NAEI, 2017 SMAI XMAI,
- 2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCACGCACTACGTG
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
2067 SMAI XMAI, 2093 DRA3,
- 2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
CCGGAGAGCGATGCGAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGGAGTGACATTGGGTC
2115 PVU2, 2159 ALWN1,
- 2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCC
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
2164 MST2, 2220 ECON1,
- 2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTAAAGACCTGGCTA
ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACCTCGTGAATTTCTGGACCGAT
2282 LysAlaLysLeuM tProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGGGGTAT
TTTCGATTGAGTACGGGTGTCGACGGACCTACGGGAAACACAGGACGGTGCGCCCATTA
2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCTCGACGTCCTTGATC
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 TGGAGTGGGACCTTCCCCATTATGCTACACACGGCCCTGTACCCCCCTTCTGCG
 ACCTCACCTGGAAGGGGTAAATACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC
 2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 CCGAATCTACAGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAGGCAG
 GGCTTGATGTGCAAGCGGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
 2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAACTCTTAATGCCCTGCCAG
 CACCCCTGAAGGTGATGCACTGCCATACCTGATGACTGTTAGAATTTACGGGCACGGTC
 2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 GTCCCATCGCCGAATTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCAGGAATACCCG
 GGGACGTTGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC
 2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 GTAGGGTCGCAATTACCTTGCGAGCCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCCTC
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCTGCACCGGCAACTGCAGGTACGAG
 2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGGGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGAGGGTATATTGTCTCTCCCGCGGCCGCTTCCAACCGCTCCCCTAGTGGG
 2850. EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnL uSerAlaProSerLeuLysAlaThrCys
 CCGCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCGGTTGAACG
 2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 ACCGCTAACCATGACTCCCTCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
 2966 ESP1, 2969 SACI,
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 GAGATGGGCGGGCAACATCACCAGGGTTGAGTCAGAAAACAAGTGGTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGAGAAATCCTG
 AAGCTAGGCGAACACCGCTCTCTCTGCTCGCCCTCTAGAGGCATGGCGCTCTTAGGAC
 3096 BGL2,
 3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 CGGAAGTCTCGGAGATTTCGCCAGGCCCTGCCCGTTTGGGCGCGGGCGGACTATAACCCC
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAAACCCGCCCGCCTGATATTGGGG
 3143 ALWN1, 3164 EAG1 XMA3,
 3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGCCCG
 GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAAGGTACCGACGGG
 3217 HGIE2, 3229 NCOI,
 3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 CTTCCACCTCCAAAGTCCCTCTCTGTGCCTCCGCTCGGAAGAAGCGGACGGTGGTCTCT
 GAAGGTGGAGGTTTCAGGGGAGGACACGAGGCGGAGCCTTCTTCGCCTGCCACCAAGGAG
 3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGTCTC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
 3332 SACI, 3346 HIND3,
 3362 SerThrSerGlyIleThrGlyAspAsnThrThrSerSerGluProAlaProSerGly
 TCAACTTCCGGCATTCGGGCGACAATACGACAACACTCTCTAGCCCGCCCTCTCTGGC
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG
 3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 TGCCCCCGCATTCGCAGCTGAGTCTATCTCTCCATGCCCCCTTGGAGGGGGAGCCT
 ACGGGGGGCTGAGGCTGCGACTCAGGATAAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
 3437 EAM11051,
 3482 GlyAspProAspLeuSerAspGlySerTrpSerTrpValSerSerGluAlaAsnAlaGlu
 GGGGATCCGGATCTTAGCGACGGGTATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG
 CCCCTAGGCCTAGAATCGTGCCAGTACCAAGTTGCCAGTCATCACTCCGGTTGCGCCTC
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,
 3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGGCGCACTCGTCACCCCGTGGCCG
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCCGGTGAGCAGTGGGGCACCGCG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
GCGGAAGAACAGAACTGCCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
CGCCTTCTTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTAA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGTATTCCACCACCTCACGCAGTGCCTTGCCAAAGGCAGAGAAAGTCACATTTGAC
AACCACATAAGGTGGTGGAGTGCCTCAGCAACGGTTCCGCTCTTCTTCAGTGTAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
AGCTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCG
TCTGACGTTCAAGACCTGTCCGTAATGGTCCTGCATGAGTTCTCTCAATTTCTGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCAACGCTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
TAGACCCACATCAACTCCGTTGGAAAGACCTTCTGGAAGACATGTAACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTACGCTGAGAAAGGGGGTTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTTCGCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACAGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCTACGGATTCCAAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACAGGACAGCGGGTTGAATTCTCTGTCGCAAGCGTGGAGTCCAAAGAAACCCCAATG
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCGCTGCTTTGACTCCACAGTCACTGAGAGCGCATCCGTACG
CCCAAGAGCATACTATGGCGCAGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,
 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
 4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCGCGTGGCCATCAAGTCC
 CTCTCCGTTAGATGTTACAACACTGGAGCTGGGGGTT^CGGGCGACCGGTAGTTCAGG
 4301 BGLI, 4308 BALI,
 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
 4322 CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGC
 GAGTGGCTCTCCGAAATACAACCCCCGGGAGAAATGGTTAAGTTCCTCCCTCTTGACGCG
 4345 APAI,
 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
 4382 TATCGCAGGTGCCCGCGAGCGCGTACTGACAACCTAGCTGTGGTAACACCTCACTTGC
 ATAGCGTCCACGCGCTCGCCGCATGACTGTTGATGACACCATTGTGGGAGTGAACG
 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
 4442 TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCGAGGGTCCAGGACTGCACCATGCTCCTG
 ATGTAGTTCCGGGCCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC
 4452 SMAI XMAI,
 CysGlyAspAspLeuValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
 4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGCGAGC
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCAGGTCTCTCGCGCGCTCG
 4508 DRD1, 4511 TTH3I,
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA
 GACTCTCGGAAGTCCGATACTGGTCCATGAGCGGGGGGACCCCTGGGGGGTGT
 ProGluTyrAspLeuGluIleThrSerCysSerSerAsnValSerValAlaHisAsp
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCACGAC
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTCACAGTCAGCGGGTGTG
 4637 SACI,
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
 4682 GGCCTGGAAAGAGGGTCTACTACCTCACCCGTGACCTACAACCCCCCTCGCGAGAGCT
 CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGAGCGCTCTCGA
 4731 NRUI,
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
 CGCACCTCTGTGTTCTGTGTGAGGTCAAGTAAAGGACCGATCCGTGTATTAGTACAAA
 AlaProThrLeuTrpAlaArgMetI1 LeuMetThrHisPhePheSerValLeuIleAla
 4802 GCCCCACACTGTGGGCGAGGATGATACTGATGCCCATTTCTTTAGCGTCTTTATAGCC
 CGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
 4806 PFLM1, 4807 DRA3,
 ArgAspGlnLeuGluGlnAlaL uAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
TCCCTGTGCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCTGCTCAGAAAAGTGGGGTACCG
TCAATGAGAGGTCCACTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
GGGAACGCTCGAAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCTT

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAGCTCAAA
CCGTCCCGACGGTATACACCGTTTATCGGAGAAGTTGACCCGTCATTCTGTGTTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTACTCCAATAGCGGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCGGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCAGCGGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTCG
TCGCCCCCTCTGTAATAGTGTGCGACAGAGTACGGGCGGGGGCGACCTAGACCAAAACG

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 CTACTCCTGCTTGTGTCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACAAACGTAACACCAACCGCGCGCCGAGGACGTCAGAGTTC
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCGCGCTCTGCAGTTCAAG

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTGAGTCGTTGGTGGAGTTTACTTGTGCCGCGAGGGGCCCTAGATTG
GGCCCCACCGCAGTCTAGCAACCACCTCAAATGAACAACGGCGCTCCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGTACCCCTTGGGCC
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCGAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGGCAATTTGGGTAAGGTCATCGAT
GGATCGACCCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGTACATACCGCTCGTCGGCGCCCTCTT
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM
GGAGGCGCTGCCAGGGCCTAATAGTCGAC
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

FIG. 22-Page 10

BEST AVAILABLE COPY

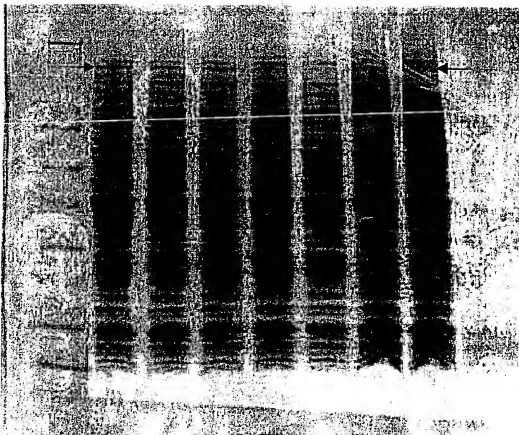


FIG. 23